# A TIME-BASED SYNCHRONOUS READOUT CMOS IMAGER

By

Qiang Luo

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2002

Copyright 2002

by

Qiang Luo

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor, Dr. John G. Harris, for his continued guidance, support and help throughout the Ph.D. graduate studies. He is not only my advisor in the academic area, but also my instructor and friend in my personal life. Without his support and encouragement, I would have taken a much longer time to finish my Ph.D. studies. I would also like to thank Dr. Robert M. Fox for his instruction and help on circuit design. Appreciation is also extended to Dr. Jose C. Principe for his instruction on sampling theory and his warm help to me during my Ph.D. studies. I would also like to acknowledge Dr. Murali Rao for his interest and participation on my supervisory committee. His support and encouragement helped me build confidence.

I also thank Dr. Zhiliang Julian Chen in Texas Instrument Inc. He proposed the TDQS system with Ravi Kishore and selflessly provided me instruction on the two-degrees of freedom time domain sampling technique and helped me start to build the system level architecture.

I also would like to thank my colleagues in the Analog Group, Ravi Kishore, Michael Erwin, Xin Qi, Xiaochuan Guo and Liping Deng, for their helpful talks with me on the proposed technique and for their friendship. Michael Erwin also developed the transform program for the conversion of ultra-wide dynamic range images.

I am especially grateful to my parents and my wife for their love and support. They provide me the strongest support and endless love during my long climb up the academic mountain. There are no words that can express my appreciation to them.

TABLE OF CONTENTS

page

ACKNOWLEDGEMENTS ..... iii

ABSTRACT..... vii

CHAPTERS

1 INTRODUCTION AND CMOS IMAGERS REVIEW ..... 1

1.1 Introduction ..... 1
1.2 Voltage Mode CMOS Imagers ..... 3
1.3 Current Mode CMOS Imagers ..... 8

2 CMOS IMAGER DEVICE PHYSICS AND TECHNOLOGY ..... 10

2.1 Overview ..... 10
2.2 Junction Photodiodes ..... 12
2.2.1 Photogenerated Current ..... 12
2.2.2 Depletion Region Width ..... 15
2.2.3 Quantum Efficiency ..... 15
2.2.4 Dark Current ..... 17
2.3 The Operation Of Photocurrent Detection ..... 19
2.4 Other Photodetectors ..... 21

3 TWO-DEGREES OF FREEDOM TIME-BASED CMOS IMAGERS ..... 23

3.1 Dynamic Range Limitations of Voltage-Based CMOS Imagers..... 23
3.2 Time-Based CMOS Imagers..... 25
3.3 Resolution Of Ultra-wide Dynamic Range CMOS Imagers..... 32
3.4 Image Reconstruction ..... 37
3.5 Examples ..... 39

4 SAMPLING STRATEGY ..... 46

4.1 Criterion Due To The Optimal SNR Consideration ..... 47
4.1.1 Sampling Strategy Without Sampling Time Limitation ..... 49
4.1.2 Sampling Strategy With A Limited Time ..... 52
4.2 Sampling Strategy For The Application Of Linear Quantization Of Light Intensity ..... 57
4.2.1 Sampling Strategy Without Sampling Time Limitation ..... 58

4.2.2 Sampling Strategy With A Limited Time ..... 62
4.2.3 Summary And Discussion ..... 63
4.3 Sampling Strategy For The Application Of Weber's Law..... 64
4.3.1 Sampling Strategy Without Sampling Time Limitation ..... 64
4.3.2 Sampling Strategy With A Limited Time ..... 67
4.3.3 Summary ..... 69
4.4 Discussion ..... 70

5 READOUT SCHEMES ..... 71

5.1 Challenges And Methods..... 71
5.2 Off-chip Memory Readout Schemes ..... 75
5.2.1 AAC Readout Scheme ..... 76
5.2.2 RAC Readout Scheme ..... 78
5.2.3 R&AAC Readout Scheme ..... 81
5.2.4 Parallel Readout Scheme..... 85
5.3 Conclusions And Discussions ..... 87

6 CIRCUIT DESIGN ..... 90

6.1 Pixel Design..... 90
6.2 Comparator Design..... 92
6.2.1 Gain Of The Preamplifier ..... 93
6.2.2 Regenerative Latch ..... 94
6.2.3 Common-Mode Input Voltage Of The Comparator ..... 96
6.2.4 Offset Of The Comparator ..... 96
6.2.5 Settling Time Of The Comparator ..... 97
6.2.6 Discussions..... 101
6.3 Disablement Circuit Design..... 102
6.4 Readout System Design..... 104
6.5 The Generation Of Reference Voltages And The "Sample" Signal ..... 107
6.6 The System Architecture Of The CMOS Imager..... 107

7 NOISE ANALYSIS ..... 109

7.1 Temporal Noise ..... 109
7.1.1 Reset Stage Noise ..... 110
7.1.2 Integration Stage Noise ..... 112
7.1.3 Readout Stage Noise ..... 113
7.1.4 Total Temporal Noise ..... 116
7.1.5 Signal To Noise Ratio And Dynamic Range ..... 116
7.2 Fixed Pattern Noise ..... 119
7.3 Defects ..... 124
7.4 Issues Of Off-chip Offset Calibration And On-chip Offset Cancellation ..... 125

8 SIMULATION AND EXPRIMENTAL RESULTS ..... 130

8.1 Single Pixel Simulation And Experimental Results ..... 130
8.2 Readout Circuit Simulation And Experimental Results ..... 134
8.3 Image Sensor Array Simulation And Experimental Results ..... 138
8.4 Discussion ..... 142

9 CONCLUSION AND FUTURE DIRECTION ..... 145

9.1 Conclusion ..... 145
9.2 Future Directions ..... 147

APPENDICES

A THE MATLAB CODE FOR THE DEMO OF IMAGE "LENA" ..... 149

B THE MATLAB CODE FOR THE DEMO OF WIDE DYNAMIC RANGE IMAGES ..... 152

REFERENCES ..... 154

BIOGRAPHICAL SKETCH..... 160

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

A TIME-BASED SYNCHRONOUS READOUT CMOS IMAGER

By

Qiang Luo

December 2002

Chairman: Dr. John G. Harris
Major Department: Electrical and Computer Engineering

A typical scene encountered in an outdoor environment where the luminance varies over a wide range (e. g. exceeding 100dB) needs an ultra-wide dynamic range imager. However, for conventional voltage-based CMOS imagers, if the integration time is too long then the brighter pixels will saturate and if the integration time is too short then the darker pixels will be swamped in noise. This voltage-based sampling technique inherently has a low dynamic range.

In this dissertation, a two-degrees of freedom time-based sampling technique is proposed for ultra-wide dynamic range CMOS imagers. The time-based approach automatically optimizes the integration time for each pixel, i.e. the bright and dark pixels choose correspondingly faster and slower integration times. It has the advantages over standard integration-mode CMOS imagers in that complex analog readout circuitry and subsequent readout noise are no longer of concern since only digital signals need to be read out from each pixel. Also, an analog-to-digital converter is no longer required and

image reconstruction is straightforward. To achieve ultra-wide dynamic range imaging in video applications, a varying voltage reference should be included to work together with the sampling time. Optimal sampling strategies are discussed in this dissertation based on maximizing the signal-to-noise ratio.

To increase the readout speed, bit plane compression readout schemes are proposed. Instead of reading out and storing the whole frame at each sampling time as conventional multiple sampling does, only the addresses of the fired pixels are read out through reconfigurable shift registers architecture.

Fixed pattern noise is evident in the captured images. Off-chip methods can calibrate somewhat the output data, but for more accurate results on-chip offset cancellation circuits are required.

# CHAPTER 1
# INTRODUCTION

## 1.1 Introduction

This dissertation studies the time-based synchronous readout CMOS imager. A typical scene encountered in an outdoor environment where the luminance varies over a wide range (e.g., exceeding 100dB) requires an ultra-wide dynamic range imager. However, for conventional voltage-based CMOS imagers, if the integration time is too long then the brighter pixels will saturate and if the integration time is too short then the darker pixels will be swamped in noise. This voltage-based sampling technique inherently has a low dynamic range. Time-based sampling technique is proposed as a way to improve the dynamic range of a CMOS imager.

The time-based approach automatically optimizes the integration time for each pixel, that is, the bright and dark pixels choose correspondingly faster and slower integration times. It has advantages over standard integration-mode CMOS imagers in that complex analog readout circuitry and subsequent readout noise are no longer of concern since only digital signals need to be read out from each pixel. Also, an analog-to-digital converter is no longer required and image reconstruction is straightforward. In time-based imagers, the pixel voltage is reset and the voltage is discharged by the photodiode current as with standard active pixel sensor imagers. However now a comparator is used to compare the pixel voltage to a voltage reference. When the threshold is passed, a digital pulse is fired. For the applications with limited sampling

time such as video applications, two-degrees of freedom, that is, the sampling times and reference voltages, are combined to improve the dynamic range. Optimal sampling strategies with the criterion of maximizing the signal-to-noise ratio are studied in this dissertation.

The time-based sampling results in multiple samples of a scene. Bit frame compression readout schemes are proposed to dramatically increase the readout speed. Instead of reading out and storing the whole frame at each sampling time as conventional multiple sampling does, only the addresses of the fired pixels are read out through a reconfigurable shift registers architecture that minimizes the readout time of pixels. The system is designed, simulated, layouted and fabricated in AMI 0.5um technology. Simulation and experimental results are presented.

Noise in the system is also analyzed. Fixed pattern noise is evident in the captured images. It mainly comes from the offset voltage variation of the comparators. Off-chip methods can calibrate somewhat the output data, but for more accurate results on-chip offset cancellation circuits are required.

In the first Chapter, the previous work in CMOS imaging is reviewed. In the second Chapter, the image sensor device physics and technology are discussed. The third Chapter presents the two-degrees of freedom time-based sampling technique and system [1]. The study on optimal sampling strategy is presented in Chapter 4. Chapter 5 shows the high-speed bit plane compression readout schemes. Chapter 6 shows the circuit design and system level architecture [2]. Noise analysis is presented in Chapter 7. Simulation and experimental results are presented in Chapter 8. Finally, conclusion and future direction are presented in Chapter 9.

Before the review of CMOS imagers, we simply classify them into classes. From the point of view of applications, there exist two classes of CMOS imagers. One is that used in general image capture, the other is that used on chip circuitry providing a processed image of a small number of features, which is named silicon retinas. From the point of view of detection method, there exist voltage-mode and current-mode CMOS imagers. Here we will first review voltage-mode CMOS imagers by dividing this large class into several small different application areas. Then, we will review the current-mode CMOS imagers.

## 1.2 Voltage Mode CMOS Imagers

In the wake of the microelectronic revolution, made possible by the astonishing advances in semiconductor fabrication technology, solid-state image sensing has developed very rapidly in the last two decades. Today there are many kinds of electronic cameras with very different characteristics [3]. The use of p-n junction devices as photodetectors can be cited as far back as the 1960s by Weckler [4]. Now, still or video cameras and camcorders have become inexpensive household items and indispensable image acquisition tools in laboratories.

High space resolution CMOS imagers are one of the important research directions. Recently, a 2048×2048 active pixel image sensor was reported by Scheffer et al. [5]. Each pixel consists of three transistors and a photodiode, with a logarithmic light-to-voltage conversion. The imager is fully random accessible in both space and time.

High-speed imaging is another important application area for CMOS imagers. Acquisition of images of fast-moving objects requires imagers with high photoresponsivity at short integration time, synchronous exposure, and high-speed

parallel readout. Stevanovic et al. [6] reported a CMOS imager with frame rate of more than 1000 frames/s, using a synchronous exposure variable between 1us to 150us. A fast synchronous shutter transistor is integrated in each pixel. Kosonocky et al. presented a multi-implant pinned-burried high-speed photodetector [7]. A 360×360 image sensor with frame rate up to $10^6$ per second was reported.

CMOS image sensors are also widely used in motion detection. A motion detector system to detect the motion of a printed random pattern on a ball used in a pointing device application was presented by Arreguit et al. [8]. It is based on the detection and tracking of spot edges passing over a matrix of pixels during the period between two light pulses. A general purpose 2-D motion detection sensor was reported [9]. It uses a binary image of zero-crossings, 2 level analog signals, the signs of spatiotemporal derivatives, 1-b multiplication and pulse widths to measure image velocity. Recently, Harrison et al. presented a Reichardt motion sensor [10]. Higgins et al. proposed a novel multi-chip neuromorphic VLSI visual motion processing system which combines analog circuitry with an asynchronous digital inter-chip communications protocol to allow more complex pixel-parallel motion processing than is possible in the focal plane [11]. Another computational sensor for active binocular tracking was reported by Lu [12].

Wide dynamic range imaging is another important application area of the CMOS imager. A typical scene encountered in nature where the luminance varies over a wide range (e.g., 120dB) needs an ultra-wide dynamic range imager. But the conventional CMOS imagers that implement voltage-based sampling schemes inherently have a small dynamic range. There have existed many different ways to combat this problem. The logarithmic response CMOS imager incorporates logarithmic compression at the

photodiode level to achieve ultra-wide dynamic range by Kavadias et al. [13]. It suffers from the problems of large fixed pattern noise (FPN) due to the device mismatches and poor sensitivity and local contrast. The multiple frame capture technique implements a lateral overflow gate to increase pixel dynamic range by Decker et al. [14]. It has the problem that FPN introduced by mismatch in the lateral overflow transistor gate-drain overlap capacitance. Also it requires capturing multiple frames and complex reconstruction processing. Yang et al.'s floating-point pixel-level ADC CMOS imagers perform well but require enormous memory to store the data and complex image reconstruction processing [15]. Recently, some implementations start to sample scenes in time domain in order to increase the imager dynamic range. W. Yang proposed a voltage-to-frequency conversion photosensor that provides high dynamic range [16], but the readout process takes very long time, and thus, it is infeasible for most real world applications. The arbitrated address event representation digital image sensor uses row and column arbiters to send out the pixels according to the firing order [17]. It requires high-resolution timer and large frame buffer. The pixel-parallel analog-to-digital (A/D) conversion CMOS imagers implement a free-running photocurrent-controlled oscillator to give a first-order $\Sigma-\Delta$ sequence [18]. A constant reference voltage is used and the imaging procedure requires as long as one second to achieve ultra-wide dynamic range.

The time-based sensing concept for CMOS imagers developed jointly between the University of Florida and Texas Instruments Inc. provides novel ideas on achieving ultra-wide dynamic range imaging for CMOS imagers. Two major methods have been proposed for reading out the data from the pixel array. One is the address event method that involves row and column arbiters to asychronously send out each pixel's address as

they fire [19]. It has the advantage of high resolution but it suffers from the problem of large errors in uniform luminance and it needs complex arbiter circuits. The other way is the bit plane readout method known as the time domain quantization sensing (TDQS) technique presented by Kishore [20], which digitizes a sensing pixel's analog value by quantizing it in the time domain. The whole bit plane is read out at each sampling instance. It uses a monotonically non-decreasing reference voltage and suffers from the problem of low resolution in high luminance range. In addition large memory is needed since each pixel is read out $2^n$ times where n is the number bits used to represent the data.

Histogram equalization is another way to increase the dynamic range of the gray levels by spreading out the gray levels of an image. It produces an increase in the average image contrast but is not always desirable. This increase of dynamic range is different from the ultra-wide dynamic range imaging that we have discussed. It is a signal processing method to enhance the captured image dynamic range. When integrating histogram equalization with CMOS imaging, we can achieve enhanced performance. Ni et al. [21] proposed a histogram-equalization-based adaptive image sensor. The image sensor can adapt automatically to different lighting conditions without frame delay and provides a constant image signal excursion range. A similar idea is proposed by Brajovic and Kanade [22].

There are many other application areas of CMOS imagers with integration of different kinds of signal processing methods. Iwata et al. proposed a functional CMOS image sensor that realizes parallel PWM readout, block averaging for gray scale image data, simple pattern detection, and calculations for x- and y- projections and centers of gravity of binary images [23]. A frame-transfer CMOS active pixel sensor was reported

by Zhou et al. [24]. The sensor architecture integrates an array of active pixels with an array of passive memory cells. Binning of pixels for variable resolution imaging is available. A CMOS imager with on-chip compression using an analog two-dimensional discrete cosine transform processor and a variable quantization level analog-to-digital converter was reported by Kawahito et al. [25]. A multiresolution CMOS imager that permits programmable multiresolution readout was presented by Kemeny et al. [26]. An on-sensor image compression for high pixel rate imaging in the video application was proposed by Aizawa et al. Conditional replenishment was implemented for compression [27-29]. A novel integration of Haar wavelet transform on CMOS imager was reported by Luo and Harris [30]. Switched capacitance circuits are implemented to realize the transform function. A simple but efficient transform unit was proposed, which is suitable for on-sensor-plane processing. Several smart image sensor arrays are presented by Schanz et al. for various applications [31]. 1-D and 2-D arrays were presented and circuit implementation complexity was discussed. Electronic shutters were also applied in CMOS imagers. Aw et al. reported a CMOS image sensor with electronic shutter integrated in each pixel, which has wide range of exposure time [32]. Iida et al. proposed a novel "I-shaped" pixel structure for high packing density devices [33]. Color imaging is also explored by researchers [34].

There also exist several special application areas of CMOS imager based on new technology and processing. Dierickx et al. proposed a very high fill factor CMOS image sensor technology by placing a small but effective electrostatic barrier between the photo sensitive volume and the unrelated junction, which is absent between the photo sensitive volume and useful junction [35]. Thin film on ASIC (TFA) technology was also applied

in CMOS imagers, which consists of an electrode (photodetector) layer, an insulation layer and an ASIC layer. Some works were reported on this application [36,37].

Seeing chips [38] or vision chips [39] are another important class of CMOS imagers. Funatsu et al. proposed a Variable Sensitivity Photo-Detector (VSPD) that realizes programmable focal plane image processing by employing between-pixel current mode calculation, together with a novel addressing method where filtering is executed by varying the addressing pattern generated by a scanning unit [40]. A retina-implant system that provides visual sensations to patients suffering from photoreceptor degeneration was reported by Schwarz et al. [41]. The CMOS imager system provides high dynamic range of more than seven decades. An ultra-high-speed smart sensory information system was proposed by Fang [42]. A neural processor is combined with CMOS APS sensor. The smart vision system on a chip can take the combined advantages of the optics and electronics.

### 1.3 Current Mode CMOS Imagers

Most CMOS imagers detect voltage signal. Some CMOS imagers detect current signal and thus are called the current-mode CMOS imagers. One of the first works on current-mode CMOS imager was reported by McIlrath et al. [43]. Noise analysis and performance limits for a current-mode active pixel sensor (APS) weres made. The imager output performance is limited by spatial fixed pattern noise. A dynamic current mirror was used to compensate threshold voltage variations. This is very necessary for future development of high-quality current-output imager. Nakamura et al. presented another current-mode active pixel sensor [44]. On-focal-plane signal processing for current-mode APS, including FPN suppression and high-resolution on-chip $\Delta$-$\Sigma$ ADC were presented.

Variable resolution CMOS current-mode APS was reported by Coulombe, et al. [45]. A current mediated APS with variable image size and resolution for power saving, electronic zooming, and data reduction at the sensor was presented.

In order to understand how semiconductor devices can be used to detect light intensity and how a CMOS imager generally works, we will discuss the basic physics and technology for CMOS imager devices in the next chapter.

# CHAPTER 2
# CMOS IMAGER DEVICE PHYSICS AND TECHNOLOGY

## 2.1 Overview

The image sensors discussed in this dissertation are primarily employed for the detection of electromagnetic radiation, especially covering the visible region (400~700 nm) of the spectrum. Here, we restrict ourselves to semiconductor image sensors based on silicon, which has been shown to be sensitive to electromagnetic radiation in the wide wavelength range between 0.1 and 1150 nm [46]. The semiconductor detectors are solid-state devices exploiting the internal photoelectric effect, as opposed to the photoemission effect of vacuum photodetectors. In them, the charge-carrier pairs generated by photon absorption contribute to electrical conduction internal to the device, and to a corresponding current at the contacts of the external circuit. The photodetection process can be schematized by the following sequence:

(1) Absorption of photons in the semiconductor with generation of charge carriers

(2) Drift of charge carriers under a suitable internal electric field

(3) Collection of the charge carriers at the ohmic contacts of the device

Since the photocurrent is typically very small (as small as a few fA to 100s of fA), it is integrated into charge and then converted to voltage before it is read out.

Figure 2.1 illustrates the energy bands of silicon at room temperature. The energy gap between the conductance band (c.b.) and valence band (v.b.) for silicon is about 1.12eV at 300°K [47]. An incident photon needs to have at least this energy to excite an

electron to jump from the v.b. to the c.b., leaving a hole behind. Since visible light wavelengths range from 400nm (violet light) to 700nm (red light), and photon energy is expressed by

$$E_{ph} = h\nu = h\frac{c}{\lambda} \tag{2.1}$$

where $h$ is Plank's constant, $c$ is the speed of light, and $\lambda$ is the wavelength, then, the visible light energy can be found to range from 2.77eV (red light) to 3.10eV (violet light). In other words, photons in the visible light range have enough energy to generate electron-hole pairs.

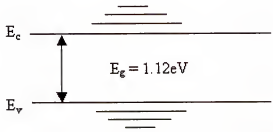



Figure 2.1 Energy band diagram of silicon. $E_g$=1.12eV at 300°K. $E_c$ is the energy of the bottom of the conductance band and $E_v$ is the energy of the top of the valence band.

In photometry incident light at a surface is denoted as luminance, which is measured in Lux (lumens/m$^2$). The values of light intensity vary over a wide range in nature. For example, Table 2.1 lists some typical values of light intensity for different environments. This wide range indicates that wide dynamic range imaging is a practical and critical issue in a single-scene capture.

Table 2.1 Typical light intensity in some environments

| environment | Light intensity (Lux) |
|---|---|
| Clear day time | $10^4$ |
| Room light | 100 |
| Full moon | 0.1 |
| Moonless night | $10^{-4}$ |

## 2.2 Junction Photodiodes

There are several types of photodetectors, the most commonly used is the photodiode, a reverse-biased pn junction. It is easy to implement in conventional CMOS processes. Also it has higher spectral response than the photogate, another widely used photodetector in CMOS imager, especially for lower wavelength (blue) due to the absorption in polysilicon gate. Photogate and other types of photodetectors are briefly discussed in section 2.4.

### 2.2.1 Photogenerated Current

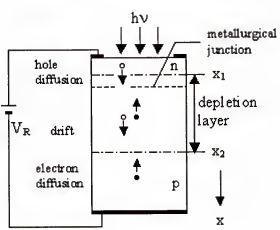



Figure 2.2 Representation of a pn junction photodiode and the photogenerated charge transport mechanism. The solid circle represents electron, and empty circle represents hole. n denotes n type material and p denotes p type material.

Figure 2.2 represents a pn junction photodiode and the photogenerated charge transport mechanism. Radiation is incident on the n-type region, which carries the ohmic cathode contact. This region is made relatively thin, by doping it more heavily than the p-region so that photons have a good chance to cross it and be absorbed in the depleted region of the junction. Assuming incident photon flux is $P_0$ photons/cm$^3$·s at the surface and the photo flux at depth $x$ is $P(x)$, then

$$\begin{cases} \dfrac{dP(x)}{dx} = -\alpha P(x) \\ P(0) = P_0 \end{cases} \tag{2.2}$$

where α is a constant related to the alleviation of photon flux. Note also that α is a function of wavelength λ. The solution of the equation is

$$P(x) = P_0 e^{-\alpha x} \tag{2.3}$$

Therefore, the electron-hole pair generation rate can be expressed as

$$G_{e_h}(x) = \frac{d(P_0 - P(x))}{dx} = \alpha P_0 e^{-\alpha x} \tag{2.4}$$

This shows that the charge carriers (electrons and holes) generated by absorption of photons with $h\nu \geq E_g$ decrease exponentially with depth $x$.

The main contribution to photogenerated current comes from the pairs generated in the depletion layer, where electron and hole are separated by the electric field internal to the junction, and by drift they reach the quasi-neutral regions where they are collected by the electrodes. This photocurrent, called drift current contribution, is associated with a fairly fast response (time scale of nanoseconds), because the speeds of electrons and holes, as shown in equation (2.5)

$$v_{n,p} = \mu_{n,p} E \tag{2.5}$$

are high. In equation (2.5), $\mu$ is the mobility, $E$ is the electric field intensity.

Conversely, photogenerated carrier pairs in the quasi-neutral regions are not equally useful because a minority carrier has a high probability of recombination with a majority carrier, thus canceling the photon detection event. The pair in itself does not generate an external current because it does not alter the charge neutrality. Only the pairs generated close to the depletion layer within a minority diffusion length (i.e., within the

mean free path traveled by the minority carriers before recombination), can contribute to collection. When the minority carrier reaches the edge of the depletion region, it is quickly swept away from the electric field and cross the junction down to the opposite region. This is called diffusion current, associated with the diffusion time constant of minority charge carriers (time scale of microseconds).

Assuming all generated carriers are collected, the drift current density can be described as

$$J_{ph} = q\int_{x_1}^{x_2} G_{e_h}(x)dx = P_0 q\left(e^{-\alpha x_1} - e^{-\alpha x_2}\right) \tag{2.7}$$

The units for $J_{ph}$ are A/cm$^2$. The diffusion current density calculation is much more complicated, and involves the diffusion constant and the recombination rate. Since it is not as useful as the drift current, we will not derive its equation here. Interested readers can find a good reference written by Gamal [48]. It can be shown that diffusion current density is also proportional to $P_0$. In summary, the following conclusions for photogenerated current can be reached:

(1) $J_{ph}$ is proportional to $P_0$, (i.e. the illumination), when the depletion layer is unchanged.

(2) A shallow pn junction ($x_1$ is small) is useful for increasing $J_{ph}$ if the depletion region has the same width.

(3) When the depletion region width ($x_2$ - $x_1$) increases, (for example, by increasing reverse bias voltage), $J_{ph}$ increases.

These properties of $J_{ph}$ are important in the implementation of photon detection, and they also provide the theoretical basis for the imaging process of CMOS image sensors.

### 2.2.2 Depletion Region Width

As we have discussed above, photocurrent is mostly composed of the drift current in the depletion layer. As a design rule, it is best to have photons absorbed in the depletion region, and thus the width of this region is made large, or at least comparable to the photon absorption length. To increase the depletion region width, it is customary to apply a reverse bias voltage to the junction. The depletion layer depths in n-type silicon and p-type silicon can be described [49]:

$$W_n = \sqrt{\frac{2\varepsilon_{Si}(\phi_0 + V_R)}{qN_D\left(1+\frac{N_D}{N_A}\right)}} \tag{2.8}$$

$$W_p = \sqrt{\frac{2\varepsilon_{Si}(\phi_0 + V_R)}{qN_A\left(1+\frac{N_A}{N_D}\right)}} \tag{2.9}$$

In these equations, $\phi_0$ is the built-in potential, and $\varepsilon_{Si}$ is the permitivity of silicon. The depletion layer width is $W = W_n + W_p$. In Figure 2.2, $W_n$ is the length from $x_1$ to the junction, and $W_p$ is the length from the junction to $x_2$. If $N_D >> N_A$, then $W_n << W_p$, in other words, the depletion layer exists almost entirely in p type silicon (the lower doped region). From equation (2.8) and (2.9), we can see that increasing the reverse bias voltage $V_R$ can widen the depletion region, but it will also increase the dark current (as discussed in section 2.2.4).

### 2.2.3 Quantum Efficiency

The quantum efficiency (QE), defined as the ratio of collected charge pairs to absorbed photons, is a function of wavelength and is represented by

$$\eta(\lambda) = \frac{\int_{x_1}^{x_2} G_{e_h}(x)dx}{P_0} = \int_{x_1}^{x_2} \alpha \cdot e^{-\alpha x} dx = e^{-\alpha \cdot x_1} - e^{-\alpha \cdot x_2} \tag{2.10}$$

This representation assumes the time scale is shorter than that of the minority carriers' diffusion time constant. If we include both the drift and diffusion current, the integration interval will be $\{x_1 - L_p, x_2 + L_n\}$, where, $L_p$ and $L_n$ are minority diffusion lengths for holes and for electrons respectively. It can be see from the equation of (2.10) that to achieve high quantum efficiency, $x_1$ should be small and $x_2$ should be big. A shallow pn junction and a wide depletion layer can satisfy this requirement.

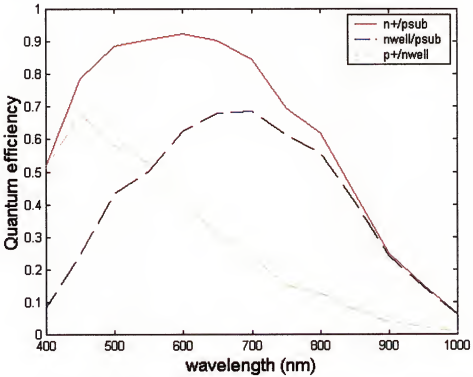



Figure 2.3 Quantum efficiency (λ=550nm) for n+/p-sub, p+/nwell and n-well/p-sub photodiodes with reverse voltage of 2.5V at temperature of 300°K.

In a standard CMOS process there are three types of photodiodes available. They are n-well/p-sub, n+/p-sub, and p+/n-well. Figure 2.3 shows how the quantum efficiency varies with photon wavelength for the three types of photodiode. It can be seen that n+/p-

sub and p+/n-well photodiodes have better QE in blue because they have smaller junction depth. n+/p-sub and n-well/p-sub photodiodes have better QE in red because they have wider depletion layer. n+/p-sub photodiode has the best overall QE among the three types of photodiodes. In reality, many process parameters affect a photodiode's QE. For example, silicide causes major QE reduction and increases the dark current (personal communication with David Yang in Pixim company, CA). It will significantly affect the QE of n+/p-sub and p+/n-well. There is no general rule and we need to examine each process individually. Being conservative, we use n-well/p-sub photodiodes exclusively in our design.

**2.2.4 Dark Current**

The current across the pn junction at a given voltage V is the sum of two contributions:

(1) The photogenerated current

(2) The usual current of the diode, given by Shockley's well-known law for the pn junction as shown in equation (2.11).

$$I_d = I_0(e^{\frac{qV}{nkT}} - 1) \tag{2.11}$$

As $I_{ph}$ is a minority current (i.e. a reverse current), it should be subtracted from the Shockley current. So, the total current is

$$I = I_0(e^{\frac{qV}{nkT}} - 1) - I_{ph} \tag{2.12}$$

where $kT/q$ = 26mV at $T$=300K, $I_0$ is the reverse saturation current, $I_{ph}$ is equal to $J_{ph}$ times junction area, and $n$ is the ideality factor of the junction. When the pn junction is reverse biased at $|V|$ >>26mV (at room temperature), (2.12) becomes

$$I = -I_0 - I_{ph} \tag{2.13}$$

That is, the photodiode supplies a current proportional to the detected incident light intensity $P_0$ superposed on a small reverse current $I_0$. Therefore, $I_0$ is also called the dark current, which implies the reverse current across the junction when there is no incident light. The sources of dark current are mainly the defects in the $SiO_2$ – Si interface, Si surface and bulk. The dark current can be represented by

$$I_0 = Aq\left[n_i^2\left(\frac{D_p}{L_p N_D} + \frac{D_n}{L_n N_A}\right) + \frac{n_i W}{2\tau}\right] \tag{2.14}$$

where $A$ is the junction area, $n_i$ is the intrinsic concentration of charge carriers, $D_p$ and $D_n$ are the minority diffusion constants (hole and electron respectively), $L_p$ and $L_n$ are hole and electron diffusion length respectively, $N_D$ and $N_A$ are doping concentrations of donors and acceptors respectively [46]. $W$ is the depletion layer width, which equates to $(W_n + W_p)$ as shown in section 2.2.2. $\tau$ is the lifetime of the charge carriers. Since

$$n_i^2 \propto T^3 \cdot e^{-\frac{E_g}{kT}} \tag{2.15}$$

$$W = f(V_R) \tag{2.16}$$

$$\tau \propto T^{-1} \cdot N_{tot} \tag{2.17}$$

It can be concluded that the dark current decreases dramatically with the decrease of temperature $T$, decreases with increase of doping concentration, and is proportional to depletion layer width. Therefore reducing the reverse bias voltage reduces the dark current, but it also reduces the photocurrent (see equation (2.7)). Generally, the typical range of the dark current at room temperature is 0.5fA - 5fA.

The existence of the dark current limits the photodetector dynamic range. Any photogenerated current close to or less than the dark current cannot be correctly measured. It also introduces unavoidable shot noise. In addition, dark current can vary widely over the image sensor array causing fixed pattern noise (FPN).

In our system design, a n-well/p-sub photodiode is applied. The reset voltage is set to be 3V to achieve a relatively wide depletion range and small dark current.

### 2.3 The Operation of Photocurrent Detection

Since the photogenerated current is very small, the most commonly used mode of photocurrent detection in an image sensor is integration, where the photocurrent is integrated over the capacitance at the photodiode cathode (called $C_{ph}$) as well as the dark current. First, the photodiode is reset to the reverse bias voltage (called $V_{res}$). Then the photocurrent and dark current discharge the $C_{ph}$ some fixed time (called the integration time $t_{\text{int}}$). At the end of the integration, the photodiode voltage $V_{ph}$ is measured and read out. Figure 2.4 shows the diagram of the integration mode for photocurrent measurement in a CMOS imager.

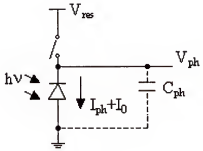



Figure 2.4 The diagram of the integration mode for photocurrent measurement.

From Figure 2.4 we can obtain

$$I_{ph} + I_0 = C_{ph} \frac{d(V_{res} - V_{ph})}{dt} \tag{2.18}$$

If the photocurrent, the dark current and $C_{ph}$ do not change with the reverse bias voltage ($V_{ph}$), the above equation shows a linear relation between $I_{ph}$ and $V_{ph}$. Consider an n-well/p-sub photodiode with $V_{res}$ = 3V and area $A$ = 123.21μm$^2$. The capacitance of photodiode $C_{diode}$ with reverse bias voltage of 3V in the AMI 0.5μm technology is (assuming an abrupt junction)

$$C_{diode} \approx \frac{40aF/\mu m^2 \cdot 123.21\mu m^2}{\left(1+\frac{3}{0.638}\right)^{0.5}} = 2.06fF \tag{2.19}$$

Taking into account the parasitical capacitance of the circuit (see chapter 5 for the circuit schematic), the total capacitance $C_{ph}$ at the cathode of the photodiode is about 17.8fF. Based on the measurements from our lab [50], we can calculate the dark current density for n-well/p-sub photodiode to be about 0.4nA/cm$^2$. Notice that depletion layer width is a function of the reverse bias voltage (here it is the $V_{ph}$) and both $I_0$ and $I_{ph}$ are functions of the depletion layer width. Thus $C_{ph}$, $I_0$ and $I_{ph}$ are all functions of $V_{ph}$. From equations (2.8), (2.9) and (2.19) we can see that these functions are all nonlinear functions. Therefore the relation of $V_{ph}$ and ($I_{ph}$ + $I_0$) is actually nonlinear and a closed form solution is impossible to obtain. Here we illustrate the simulation results in Figure 2.5. The relation between $V_{ph}$ and $I_{ph}$ that takes into account the effect of $V_{ph}$ fits the linear model very well under low luminance. However the error increases for high luminance cases and it becomes worse as the integration time $t_{\text{int}}$ increases. The error causes an overestimation of the actual signal since the signal equals to ($V_{res}$–$V_{ph}$). Figure 2.6 and 2.7 show the dependence of $C_{ph}$ on $V_{ph}$ and $I_{ph}$ on $V_{ph}$ during the integration respectively. There is about a 2% variation in $C_{ph}$ and 0.27% variation in $I_{ph}$ for 16lux light luminance; about 17.85% variation in $C_{ph}$ and 3.05% variation in $I_{ph}$ for 80lux light luminance. The

major non-linearity comes from $C_{ph}$ because $I_0$ is generally much less than $I_{ph}$. It is also clear that the error becomes greater as $V_{ph}$ drops closer to 0. This situation generally occurs for high luminance pixels that are integrated for sufficient long time. It brings nonlinear errors for conventional voltage-based CMOS imagers (see chapter 3 for the concept) because all the pixels are integrated for the same time (e.g., 30ms for video application). In time-based CMOS imagers (see Chapter 3 for the concept) the high luminance pixels have less integration time than the low luminance pixels. For example, the 16lux light is integrated for 8ms and the 80lux light is integrated for 4ms. Therefore, the non-linearity does not affect time-based imagers as much (see Figure 2.5) and we can ignore the nonlinear errors.

## 2.4 Other Photodetectors

There exist many other types of photodetectors, such as PIN junction photodetectors, avalanche photodiodes, and photogates. A photogate is sometimes used in image sensors. For example, in frame transfer charge-coupled devices (CCDs) and CMOS photogate active pixel sensor (APS). In CMOS technology photogate is implemented using an MOS gate and drain. Here we will not discuss photogate or other photodetectors in detail. Interested readers can refer to the literature [46,48,51,52].

In the next chapter, we will propose the concept of two-degrees of freedom time-based CMOS imager and discuss how a time-based CMOS imager can extend the dynamic range.

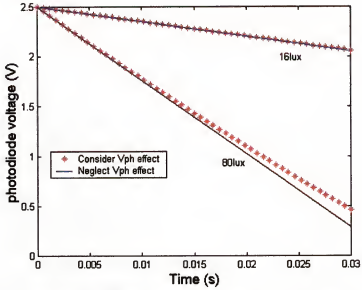



Figure 2.5 Comparison of linear model described by (2.18) and the actual model taking into account the effect of $V_{ph}$.

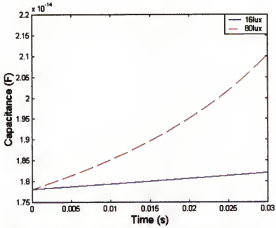



Figure 2.6 The capacitance change due to the change $V_{ph}$ change in the integration procedure.

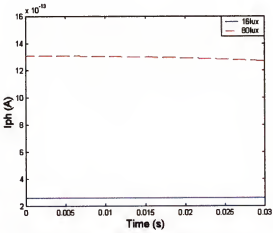



Figure 2.7 The photogenerated current due to the $V_{ph}$ change in the integration procedure.

# CHAPTER 3
# TWO-DEGREES OF FREEDOM TIME-BASED CMOS IMAGERS

Dynamic range (DR) is one of the most important features of CMOS imagers. Assume voltage is measured as the representation of signal. Dynamic range is defined as the ratio of maximal correctly measurable signal to the noise floor (equation (3.1)).

$$DR = \frac{V_{\max}}{v_{noise_floor}} \tag{3.1}$$

Either increasing the maximal detectable signal level or decreasing the noise floor can increase the dynamic range. The noise floor is decided by the physics of devices and semiconductor processes and is therefore hard to reduce. The maximal correctly measurable signal is typically limited by the power supply voltage, which is slowly creeping downwards with each new generation of CMOS processes. Here we will focus on increasing the maximal detectable signal to increase the dynamic range without the constraint of power supply voltage. To do this, a two-degrees of freedom time-based CMOS imager is proposed. Before discuss the details, we will briefly review the dynamic range limitations of conventional CMOS imagers that measure the photodiode voltages. These conventional CMOS imagers are called voltage-based CMOS imagers.

## 3.1 Dynamic Range Limitations Of Voltage-Based CMOS Imagers

We have reviewed different voltage-based CMOS imagers in Chapter 1. In general, these CMOS imagers measure the photodiode voltage $V_{ph}$ after a certain integration time $t_{\text{int}}$. A typical three-transistor active pixel sensor (APS) CMOS imager

pixel structure is shown in Figure 3.1. The function of the APS sensor is represented by the following equation

$$\Delta V = V_{res} - V_{ph} = \frac{I_{ph} \cdot t_{int}}{C_{ph}} \tag{3.2}$$

where $V_{res}$ is the reset voltage, $I_{ph}$ is the photocurrent and $C_{ph}$ denotes the capacitance at photodiode cathode.

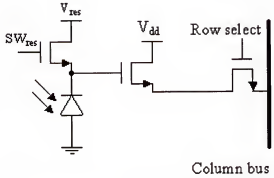



Figure 3.1 The 3 transistors structure of the pixel of an APS CMOS imager.

It is clear that $V_{ph}$ decreases linearly with integration time with a slope of $-\frac{I_{ph}}{C_{ph}}$.

This function is shown in Figure 3.2.

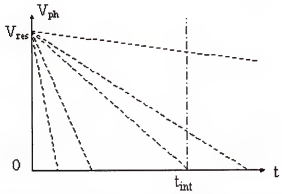



Figure 3.2 Voltage-based CMOS imagers working scheme.

All the pixels integrate the photogenerated current for a fixed time $t_{int}$. At the end of the integration, the photodiode voltages are read out. For those pixels with high luminance, the photogenerated currents are so large that the photodiode voltages drop to 0 before the end of the integration. This is also called pixel saturation. Therefore, the values of these pixels cannot be correctly measured. The maximal detectable photocurrent is decided by

$$I_{ph,\max} = \frac{C_{ph} \cdot V_{res}}{t_{\text{int}}} \tag{3.3}$$

In Figure 3.2, the third dashed line indicates the case of the maximal detectable photocurrent. On the other hand, the photodiode voltages drop is very small for those pixels with low luminance that the change of voltages may not be able to be detected. In general, the minimal correctly measurable photocurrent is about the dark current. Here, we give an example of the typical value of dynamic range of a voltage-based CMOS imager. Assume $C_{ph}$ equals to 5fF, $V_{res}$ is 2.5V, and $t_{\text{int}}$ is 30ms, we can obtain $I_{ph,max}$=0.42pA from equation (3.3). Assume the minimal correctly measurable photocurrent is 0.5fA, the dynamic range of this imager is 58.5dB. Since $V_{res}$ cannot be infinitly large, the maximal correctly measurable photocurrent is limited by the maximal available $V_{res}$ in circuits.

In summary, the linear relationship between $V_{ph}$ and $I_{ph}$ functionally causes bright pixels to saturate, resulting in a relatively small dynamic range.

### 3.2 Time-Based CMOS Imagers

In order to improve the dynamic range of a CMOS imager, we have to break through the barrier of voltage-based measurement, duo to the linear relationship between $V_{ph}$ and $I_{ph}$. Figure 3.3 depicts another idea of photomeasurement for a CMOS imager,

where time is measured to detect the incident light intensity. Compared with the conventional voltage-based CMOS imagers, the CMOS imagers involving this method are called time-based CMOS imagers. The basic mechanism of the time-based CMOS imagers is to measure the integration time that a pixel photodiode voltage needs to drop to a certain reference voltage.

The simplest time-based CMOS imager is to set the reference voltage as a constant value and measure the integration time for each pixel to drop from the reset voltage to the reference voltage. Several researchers have already implemented this method [17-18]. Figure 3.3(a) illustrates the working scheme. It is clear that those high luminance pixels that saturate and cannot be correctly measured in the voltage-based CMOS imagers (see Figure 3.2), are now able to be correctly measured with a time-based strategy by reading out the time needed for the pixel photodiode voltages to drop to $V_{ref}$. If the minimal correctly measurable photocurrent is the same as that in the voltage-based CMOS imagers, the dynamic range of the time-based CMOS imagers becomes wider than the voltage-based CMOS imagers, but it may take a long time for $V_{ph}$ of the low luminance pixels to drop to the reference voltage.

As we discussed above that with a constant reference voltage, the extra dynamic range for reasonable array sizes is achieved at the expense of very long integration times (up to seconds). These long integration times are not appropriate for dynamic scenes and cannot be used for video mode imaging. Interestingly, if video mode imaging is used with a maximum integration time of 33ms, current CMOS processes limit these integrate and fire imagers to around 60dB of dynamic range which is about the same as the conventional CMOS imagers described above. Another type of time-based CMOS imager

is the sorting image sensor demonstrated by Brajovic et al. [22]. The sensor chip detects an image focused on its sensitive area and computes an image of indices. During the process, the chip generates a waveform that represents a cumulative histogram of the detected scene and uses this waveform to map detected light intensities into indices. This waveform can be used for mapping indices back to the received intensities. The sorting sensor performs a global operation over an entire imaging array. These previous time-based imagers are unable to take advantage of a time-varying reference voltage.

Other approaches have been studied to improve the dynamic range by incorporating the use of reference, in one form or another, as a control variable. Ni et al. demonstrated a histogram-equalization-based adaptive image sensor using an increasing ramp which exceeds the precharged voltage of the photodiodes after the limited exposure time [21]. In the locally autoadaptive image sensor LAR II developed by Silicon Vision GmbH [36], the integration time for each pixel is determined by how long it takes the sense node voltage to equal or pass a given reference level. The integration time is measured in intervals and a voltage ramp that ramps up one step after every comparison is used as timestamp information to record the number of intervals required to cross the reference level. Astrom et al. [53] adjusted reference level to realize linear mapping between the time to reach the reference level and the intensity level in the application of near-sensor image processing. The reference level turns out to be a second-order polynomial with a parabola shape. The time interval must be larger than the time consumed for program loop. Although this method is not related to improving dynamic range of an imager, it is one of the first works on how to select a varying reference level based on application requirements.

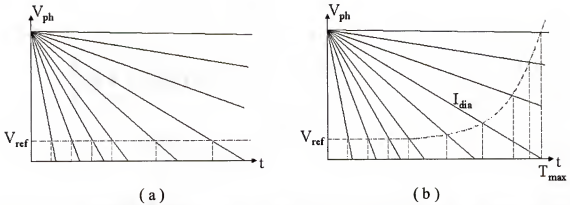



Figure 3.3 The working schemes of time-based CMOS imagers. The time instance when $V_{ph}$ drops to $V_{ref}$ is measured. (a) $V_{ref}$ has a constant value. (b) $V_{ref}$ is a function of time.

For video applications, the integration time must be less than about 33ms (frame rate of 30Hz). An option is to still use a constant reference voltage. However, those photocurrents having a magnitude less than $I_{dia}$ (see Figure 3.3(b)) will never trigger the comparator. At $T_{\max}$, the analog voltages could be read out but this requires an ADC and other analog processing as conventional CMOS imagers. This solution is not preferred since the advantages of time-based CMOS imagers cannot be totally realized and this mixed-signal processing results in a larger pixel size (two sets of processing circuits) and much more complicated sampling and reconstruction [54]. If no analog readout is implemented, all the $I_{ph}$'s less than $I_{dia}$ will be assigned the value of $I_{dia}$ and thus, the dynamic range is dramatically reduced. Another idea is to enable every pixel to trigger in the limited time by setting the reference voltage to be slightly below $V_{res}$. However, this approach has two major problems. First, well charge is badly used and second the sampling interval for high light intensity is too smaller to be feasible. Thus, to achieve an ultra-wide dynamic range and meanwhile satisfy the finite integration time requirement, a varying reference voltage is preferred over any constant reference voltage strategy. Figure 3.3(b) illustrates this idea. The low luminance pixels can now be measured in a

much shorter integration time. The combination of the two-degree of freedom, sampling times and time-varying reference voltages, provides the ultra-wide dynamic range imaging even in video applications.

Let's explain the change further. Equation (3.3) can be rewritten as

$$f_{\text{int}} = \frac{1}{t_{\text{int}}} = \frac{I_{ph}}{C_{ph} \cdot \Delta V} \tag{3.4}$$

An image with a dynamic range of 120dB indicates an $I_{ph}$ of 120dB dynamic range. While the dynamic range of neither the $f_{\text{int}}$ nor the $\Delta V$ can have so high a value. For example, in video applications, the integration time must be less than 33ms, but typically the minimal integration time $t_{\text{int,min}}$ will be a few microseconds. So the dynamic range of $t_{\text{int}}$ or $f_{\text{int}}$ is about 60dB. For $\Delta V$, the maximum value is $V_{dd}$ (e.g., 5V), and the minimal value is about a few millivolts due to the offsets in the circuit. Thus, the dynamic range of $\Delta V$ is also about 60dB. It is clear that neither of them can represent the 120dB dynamic range input light singly. Since voltage-based CMOS imagers only measure the $\Delta V$, they cannot provide a dynamic range as high as 120dB. Similarly, those time-based CMOS imagers with a constant reference voltage only measure the $t_{\text{int}}$ and therefore cannot provide a dynamic range as high as 120dB either. How can we break through the limitations of dynamic range? Notice that

$$\lg I_{ph} = \lg C_{ph} + \lg f_{\text{int}} + \lg \Delta V \tag{3.5}$$

where $I_{ph}$ has a range of about 0 to 120dB, $f_{\text{int}}$ has a range of about 0 to 60dB in video applications, $\Delta V$ also has a range of about 0 to 60dB, and $C_{ph}$ is a constant. It is clear that only an integration time change or only a $\Delta V$ change cannot satisfy the equation (3.5). But if both terms are variable and contribute about 60dB, equation (3.5) then can be

satisfied. This will break through the conventional CMOS imager limitations. Figure 3.3(b) illustrates this idea.

Furthermore, equation (3.6) shows that the ratio of two pixel intensities is the ratio of integration time and $\Delta V$ products.

$$\frac{I_{ph2}}{I_{ph1}} = \frac{t_{\text{int}1} \cdot \Delta V_2}{t_{\text{int}2} \cdot \Delta V_1} \tag{3.6}$$

If $I_{ph1}$ is $10^6$ times of $I_{ph2}$, we only need to make $t_{\text{int}2}\Delta V_1$ be $10^6$ times of $t_{\text{int}1}\Delta V_2$. Combined with equation (3.5), it is clear that as long as $\lg f_{\text{int}} + \lg \Delta V \geq 120dB$ in equation (3.5), we can achieve this ultra-wide dynamic range without requiring both $t_{\text{int}}$ and $\Delta V$ to have the dynamic range of about 60dB. For instance, in some applications such as still imaging, the maximal integration time is longer than 33ms and has a dynamic range of greater than 60dB but the reference voltage has a smaller dynamic range than 60dB. In conclusion, the two degrees of freedom provide a lot of flexibility in achieving ultra-wide dynamic range imaging. Next we will describe its working mechanism.

In the time-based CMOS imagers, the active pixel sensors (APS) that are typically used in the voltage-based CMOS imagers, are no longer suitable. Instead, a digital pixel sensor (DPS) must be used (Figure 3.4).

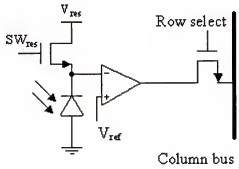



Figure 3.4 A typical DPS pixel schematic.

The basic mechanism of a DPS pixel is: the incident light discharges the reset photodiode, and the voltage across the photodiode ($V_{ph}$) is compared with the reference voltage ($V_{ref}$). Once $V_{ph}$ drops below $V_{ref}$, the comparator toggles and the digital output status changes. The measured state-change time instance and the reference voltage will be used together to calculate the photocurrent and thus, reconstruct the image. Figure 3.5 shows the mechanism of the sampling and the reconstruction of an image for a time-based CMOS imager. $V_{ref}$ is applied to all the pixels so that bright pixels always toggle the comparator before dark pixels and causing the sampling order from bright pixels to dark pixels. In general, $V_{ref}$ is small for bright pixels and large for dark pixels. For example, Kishore implemented a non-decreasing $V_{ref}$. The "DN" in Figure 3.5 denotes "digital number". In the sampling mode, the digital numbers that represent the sampling order (index) and the pixel positions are read out at each sampling time instance. In the reconstruction mode, these indices, together with the $V_{ref}$, are used to calculate the $I_{ph}$ to reconstruct the image. Note that a $V_{ref}$ change will not affect $V_{ph}$.

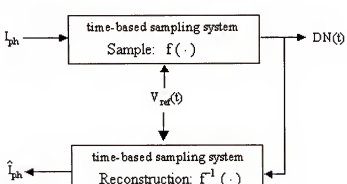



Figure 3.5 The diagram of the sampling and reconstruction in time-based CMOS imagers.

Here, let us use an example to examine the idea of achieving ultra-wide dynamic range using both integration time and reference voltage. Without loss of generality,

assume the capacitance at the photodiode cathode is 5fF or $C_{ph} = 5fF$, $I_{ph,\max} = 0.5nA$, $I_{ph,\min} = 0.5fA$. Also assume the maximum integration time is $t_{\text{int,max}} = 33ms$ and the minimum integration time is set to be $t_{\text{int,min}} = 10\mu s$. Assume $V_{ref}$ has the range of 2mV to 2.5V, and $V_{res} = 2.5V$. To catch the brightest incident light,

$$t_{\text{int,min}} = \frac{C_{ph} \cdot (V_{res} - V_{ref})}{I_{ph,\max}} = 10\mu s$$

This requires that $V_{ref} = 1.5V$. Similarly, after the integration of $t_{\text{int,max}}$, the smallest photocurrent should be measured. This requires that $V_{ref} = 2.4967V$ or just 3.3mV below the reset voltage. Therefore, the integration time contributes 70.4dB dynamic range ($\frac{t_{\text{int,max}}}{t_{\text{int,min}}}$) and the reference voltage contributes 49.6dB dynamic range ($\frac{1.0V}{3.3mV}$, here assuming non-decreasing reference voltage) to the whole 120dB dynamic range of incident luminance ($\frac{I_{ph,\max}}{I_{ph,\min}}$).

### 3.3 Sampling Resolution of Ultra-Wide Dynamic Range CMOS Imagers

As discussed above, an ultra-wide dynamic range can be achieved using the time-based sampling scheme. Since the time-based sampling scheme provides a nonlinear relation between sampling time instant and photocurrent, the next problem is how to achieve a reasonable resolution for the time-based sampling. This problem becomes outstanding in the high luminance range because the photodiodes discharge very fast but the integration interval cannot be infinitely small. On the other hand, it is much easier to achieve high resolution in the low luminance range. Figure 3.6(a) illustrates the relation between $I_{\text{ph}}$ and $t_{\text{int}}$ when $\Delta V$ is a constant. It is clear that a uniform sampling in this

scheme may result in under-sampling in the high luminance range and over-sampling in the low luminance range due to the non-linearity. Therefore, nonuniform sampling schemes may need to be adopted. To improve the resolution in the high luminance range, we hope to increase the slope of the $I_{ph}$~$t_{int}$ curve, and to avoid the over-sampling in the low luminance range, we hope to decrease the slope of the $I_{ph}$~$t_{int}$ curve (see Figure 3.6(b)). In order to realize this modification, a $V_{ref}$ change must be involved and therefore, based on the two degrees of freedom ($t_{int}$ and $V_{ref}$) system, a reasonable sampling resolution can be achieved.

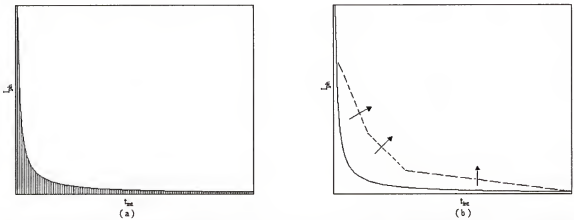



Figure 3.6 (a) The nonlinear relation between $I_{ph}$ and $t_{int}$ for a constant $V_{ref}$. A uniform sampling results in under-sampling in the high luminance range and over-sampling in the low luminance range. (b) The dash lines represent a possible relation between $I_{ph}$ and $t_{int}$ for a varying $V_{ref}$. They help provide better resolution in high luminance range and avoid over-sampling in low luminance range. The arrows represent the directions of changes.

Certainly the goals of the application must be considered before choosing the sampling scheme. For instance, if the image will only be viewed by humans on a monitor that only has a dynamic range of about 8 bits, it is not necessary to represent the full $2^{20}$ digital resolution for a 120dB dynamic range image. In fact, $2^8$ digital numbers need to be assigned so that these 256 points can cover the whole dynamic range. Thus, a reasonable distance between two sampled light intensities should be taken into account. There exist many ways to sample the ultra-wide dynamic range luminance based on the idea as

shown in Figure 3.6(b). First, let's analyze what are the factors that are related with the selection of resolution and how they work.

Equation (3.6) is shown below for convenience.

$$\frac{I_{ph2}}{I_{ph1}} = \frac{t_{\text{int}1} \cdot \Delta V_2}{t_{\text{int}2} \cdot \Delta V_1}$$

If $I_{ph2} = I_{ph1} - \Delta I_{ph}$ and $t_{\text{int}2} = t_{\text{int}1} + \Delta t_{\text{int}}$, here $\Delta I_{ph}$ and $\Delta t_{\text{int}}$ are functions of time. From this equation, we can obtain

$$\left(\frac{\Delta V_2}{\Delta V_1}\right) = \left(1 - \frac{\Delta I_{ph}}{I_{ph1}}\right)\left(1 + \frac{\Delta t_{\text{int}}}{t_{\text{int}1}}\right) \tag{3.7}$$

Further more

$$\left(\frac{\Delta V_2}{\Delta V_{brightest_pixel}}\right) = \left(1 - \frac{\Delta I_{ph}}{I_{ph,\max}}\right)\left(1 + \frac{\Delta t_{\text{int}}}{t_{\text{int},\min}}\right) \tag{3.8}$$

Clearly $\Delta V_2$ is a function of $\Delta I_{ph}$ and $\Delta t_{\text{int}}$. The required resolution of $\Delta I_{ph}$ and the integration time provide the constraints for deciding the reference voltage. Let us see an example on the sampling pattern with the consideration of resolution.

Assume 10-bit sampling is used to capture a 20 bits dynamic range image. The first example used here is to illustrate that a quadratic change of $V_{\text{ref}}$ can be implemented to improve the resolution in the high luminance range instead of a non-decreasing $V_{\text{ref}}$. Assume $C_{ph} = 5fF$, $I_{ph,\max} = 0.5nA$, and $t_{\text{int},\min} = 10\mu s$, the $\Delta V$ of $I_{ph,max}$ is $\Delta V = \frac{I_{pn,\max} \cdot t_{\text{int},\min}}{C_{ph}} = 1V$. Assume the minimal allowable sampling interval is 5μs, the following results (Table 3.1) will be achieved if $\Delta V$ is kept as a constant of 1V.

Table 3.1 An example with constant reference voltage

| $t_{int}$(us) | ΔV(V) | Normalized $\frac{I_{ph}}{I_{ph,max}}$ |
|---|---|---|
| 10 | 1 | 1 |
| 15 | 1 | 0.67 |
| 20 | 1 | 0.5 |
| 25 | 1 | 0.4 |
| 30 | 1 | 0.33 |
| 35 | 1 | 0.28 |
| 40 | 1 | 0.25 |
| 45 | 1 | 0.22 |
| 50 | 1 | 0.2 |
| … | … | … |

In the above case, the resolution in the high end of dynamic range of between normalized $I_{ph}$ of 1 and 0.5 is small. There is only 1 sample point. If $\Delta V$ is monotonically decreasing, a worse resolution will be achieved. How can we improve the resolution? For example, can we capture the photocurrents of 1, 0.9, 0.8, 0.7, 0.6 and 0.5? From equation (3.8), we can calculate the corresponding $\Delta V$. Here a quadratic change on $\Delta V$ can help to achieve the required resolution. This is shown in table 3.2.

Table 3.2 An example with varying reference voltage

| $t_{int}$(us) | ΔV(V) | Normalized $\frac{I_{ph}}{I_{ph,max}}$ |
|---|---|---|
| 10 | 1 | 1 |
| 15 | 1.35 | 0.9 |
| 20 | 1.6 | 0.8 |
| 25 | 1.75 | 0.7 |
| 30 | 1.8 | 0.6 |
| 35 | 1.75 | 0.5 |
| 40 | 1.6 | 0.4 |
| 45 | 1.35 | 0.3 |
| 50 | 1 | 0.2 |
| … | … | … |

Here assume ΔV=1.8V is available.

More gradations in the bright end of the dynamic range are achieved. Of course, we cannot achieve too high resolution in this high luminance range due to the dynamic

range limit of $V_{\text{ref}}$. When we go to the lower luminance ranges, a non-decreasing $V_{\text{ref}}$ may be enough to achieve the required resolution. Let's continue the example and generate the rest of Table 3.2. Assume it is decided to use 95 points to cover the range of $0.01I_{ph,\max}$ to $0.2I_{ph,\max}$, 450 points to cover the range of $0.001\ I_{ph,\max}$ to $0.01I_{ph,\max}$, 450 points to cover the range of $10^{-4}I_{ph,\max}$ to $10^{-3}I_{ph,\max}$, 10 points to cover the range of $10^{-5}I_{ph,\max}$ to $10^{-4}I_{ph,\max}$, and 10 points to cover the range of $10^{-6}I_{ph,\max}$ to $10^{-5}I_{ph,\max}$. This sampling scheme samples more points in the middle band of the whole dynamic range but fewer point in the high and low ends. Table 3.3 shows a possible way to achieve this sampling.

The $\Delta V$ change versus time is plotted in Figure 3.7. It shows a quadratic change of $\Delta V$ at the very high luminance range, a constant $\Delta V$ at the middle luminance range and a monotonically decreasing $\Delta V$ at the low luminance range. Figure 3.8 illustrates the $f_{\text{int}}$ versus $I_{ph}$ in the sampling mode. Clearly the dynamic range is compressed in the low luminance range and partly compressed in the very high luminance range. The total dynamic range of $f_{\text{int}}$ is about 69.5dB. Note that of course for this example there exist other ways to realize the sampling from 10us to 30ms.

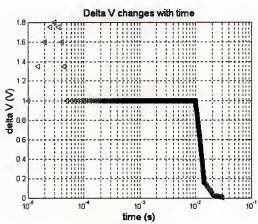



Figure 3.7 ΔV versus time in the example.

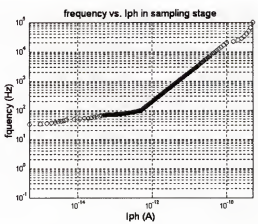



Figure 3.8 $f_{\text{int}}$ versus $I_{\text{ph}}$ in the example.

Table 3.3 An example of sampling strategy

| $t_{int}$(us) | ΔV(V) | Normalized $\frac{I_{ph}}{I_{ph,\max}}$ |
|---|---|---|
| 60us | 1 | 0.166667 |
| 70us | 1 | 0.142857 |
| 80us | 1 | 0.125000 |
| 90us | 1 | 0.111111 |
| 100us | 1 | 0.100000 |
| … | … | … |
| 1000us | 1 | 0.010000 |
| 1020us | 1 | 0.009804 |
| 1040us | 1 | 0.009615 |
| … | … | … |
| 9960us | 1 | 0.001004 |
| 9980us | 1 | 0.001002 |
| 10000us | 1 | 0.001000 |
| 10010us | 1-1.9m | 0.000997 |
| 10020us | 1-3.8m | 0.000994 |
| … | … | … |
| 14500us | 145.0m | 0.000100 |
| 15.15ms | 132.6m | 0.000090 |
| 15.8ms | 120.2m | 0.000080 |
| … | … | … |
| 21ms | 21.0m | 0.000010 |
| 22ms | 19.8m | 0.000009 |
| 23m | 18.4m | 0.000008 |
| 24ms | 16.8m | 0.000007 |
| 25ms | 15.0m | 0.000006 |
| 26ms | 13.0m | 0.000005 |
| 27ms | 10.8m | 0.000004 |
| 28ms | 8.4m | 0.000003 |
| 29ms | 5.8m | 0.000002 |
| 30ms | 3.0m | 0.000001 |
| | | 0 |

### 3.4 Image Reconstruction

Knowing the sampling time instance $t_{int}(n)$ and the corresponding reference $V_{ref}(n)$, we can reconstruct the image using

$$\hat{I}_{ph}(n) = \frac{C_{ph}\left(V_{res} - V_{ref}(n)\right)}{t_{int}(n)} \tag{3.9}$$

where $n$ (= 1, 2, ...) represents the sampling index. From the above discussion we know that the sampled data is a nonlinear representation of the original image. This can be clearly seen in Figure 3.6. The $t_{\text{int}}$ or $f_{\text{int}}$ is a nonlinear mapping of the original image, but a reconstructed image should be the linear mapping of the original one. In order to change this nonlinear relation to a linear relation, $\Delta V$ should be a constant value (see equation (3.4)). Since $\Delta V$ is a linear function of integration time for a known luminance, we just need to choose a $\Delta V$ as a reference and convert other $\Delta V$'s to this reference value by multiplying them with a proper factor. If the $\Delta V$ used for $I_{\text{ph,max}}$ (named as $\Delta V_1$) is chosen to be the reference value, then the revised $f_{\text{int}}$ (named as $f_{\text{int}}'$) is

$$f'_{\text{int},i} = f_{\text{int},i} \cdot \frac{\Delta V_i}{\Delta V_1}, \qquad (i = 2,3,4,...) \tag{3.10}$$

Figure 3.9 depicts the $f_{\text{int}}'$ versus $I_{\text{ph}}$ for the example in section 3.3. Clearly that a linear relation is achieved between $f_{\text{int}}'$ and $I_{\text{ph}}$, and $f_{\text{int}}'$ has a dynamic range of 120dB. Thus, $f_{\text{int}}'$ is a linear mapping of $I_{ph}$ and can be used as a scaled reconstructed image.

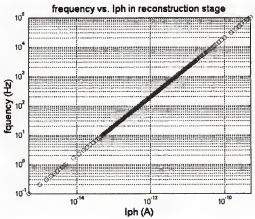



Figure 3.9 $f_{\text{int}}'$ versus $I_{ph}$ for the example in section 3.3.

We need to note that although $f_{\text{int}}'$ represents the reconstructed image, it may have a too wide dynamic range (e.g. about 120dB) to be displayed fully on a small dynamic

range environment (e.g. 50dB CRT). On the other hand, the $f_{int}$ represents the sampled image that is a nonlinear mapping (compressed dynamic range) of the original image and can be fully displayed on a CRT.

### 3.5 Examples

Here we will show some more examples to present the two-degrees of freedom time-based sampling technique. The first example is to use 6 bits per sampling to represent an 8-bit dynamic range image. The 8-bit dynamic range image used here is a 256 gray scales image of "lena". The resolution of the image is 256×256. In order to present the ability of capturing the details in the high luminance range, we added a small bright window at the left top corner of the image with some bright stripes inside (see Figure 3.10(a), 256 gray scales display). Assume we only have a 6-bit display environment, a normalized $C_{ph}$=1, and the sampling interval of $\Delta T = 1$. The sampled 6-bit dynamic range image is shown in Figure 3.10(b) with 64 gray scales display. We can see that those details in the bright window are captured. The sampled image not only captured the entire dynamic range of the original image, but also had some resolution in the high luminance range.

Figure 3.10(c) (64 gray scales display) illustrates the reconstructed image based on equation (3.10). It is also displayed in the 6-bit display environment. A lot of details are lost compared with the image in Figure 3.10(b) because the pixel values are out of the display range. To verify the captured image has a wide dynamic range, the pixel values of Figure 3.10(c) are divided by 2 each time and those details "hidden" in the bright background appear gradually. Figure 3.10(d) (64 gray scales display) shows the result that the pixel intensities of Figure 3.10(c) are divided by 4. The information is not lost,

but it is not visible in the display. The nonlinear display (as shown in Figure 3.10(b)) is suitable to depict all the information. A demo program has been developed to show the sampled image and reconstructed image (see appendix A).

The sampling time/order versus normalized photocurrent is depicted in Figure 3.11. The $\Delta V$ change makes the solid line curve piece-wise linear. The change strategy is like that shown in Figure 3.6(b).

The second example is a set of ultra-wide dynamic range images provided in Debevec [53]. The original images are shown in the left column of Figure 3.12. These are 97~168dB dynamic range images with the size of 480×720 or 768×512. A log luminance representation is adopted here. This representation means that at any exposure level, there will be equal brightness steps between values. This corresponds well with human visual response, whose contrast threshold is constant over a wide range of adaptation luminance (Weber's law) as illustrated in Larson [54]. By separating luminance and using a log encoding, we can use 8 bits to record nearly 5 to 6 orders of magnitude in 5% relative steps that will be imperceivable under most conditions. This relation is obtained by

$$10^6 \cdot r^{256} = 1 \tag{3.11}$$

where $r$ denotes the relative brightness step. This equation gives that $r = 94.75\%$, or $(1-r) = 5.25\%$. It is clear that if more bits are used to record an image, the relative step will become smaller. For example, using 10 bits to record an image in log luminance representation gives $(1-r) = 1.34\%$. This is good for high quality imaging. Here, we use an 8-bit log luminance representation as an example. As we discussed before, the resolution in the very high luminance range is limited by the minimal sampling interval. Here a typical value of 5us is adopted as the minimal sampling interval and $V_{res}$ is set to

be 2.5V. It is found that a 5% relative step is too small step to be realized in the very high luminance range. Thus, the minimal relative step for very high luminance range is adopted after experiments, whose results give a 10% relative step for $0.2 \cdot I_{ph,\max}$ to $I_{ph,\max}$ range. Sixteen samples can be assigned to this range.

The images in the right column of Figure 12 illustrate the sampled results. It is clear that the ultra-wide dynamic range original images are nicely captured and displayed with 256 gray scales. Those invisible details in original images are now visible in captured index images with good quality. A demo program will show the sampled image and the reconstructed image. Codes can be found in appendix B.

In the last two sections, we have given three different sampling pattern examples of the two degrees of freedom time-based CMOS imagers. It is worth to repeat that this system is very flexible. In a real application, a set of sampling patterns can be built in the system and selected by customers according to the application constraints. As we have reviewed in the beginning of this Chapter, there exist various works on selecting time-varying reference voltage to help extend dynamic range. However, all these previous works do not provide a way to optimally select the time-varying reference voltages and there is no instruction on how to find an optimal sampling strategy from all the possible ones. In the next chapter, we will explore the problem of how to pick a sampling strategy. We will discuss what is an optimal sampling strategy for maximizing the signal-to-noise ratio and how to realize it.

Figure 3.10 Example of figure "lena". (a) The original 8-bit dynamic range image is displayed with 256 gray scales. (b) The sampled image with only 6 bits per sample is displayed with 64 gray scales. (c) The reconstructed image. The image has 48dB dynamic range and is displayed with 64 gray scales. Most details cannot be displayed due to the limit display dynamic range. (d) The image with all pixel intensities being one quarter of those in (c). This image is also displayed with 64 gray scales. The invisible details in (c) can now be seen. This example shows that all the details are captured in (c) but are just invisible due to the finite display dynamic range.

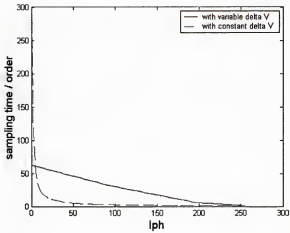



Figure 3.11 The sampling time/order versus normalized photocurrent. The red dashed line is the result for constant $V_{ref}$ and the blue solid line is the result for varying $V_{ref}$.

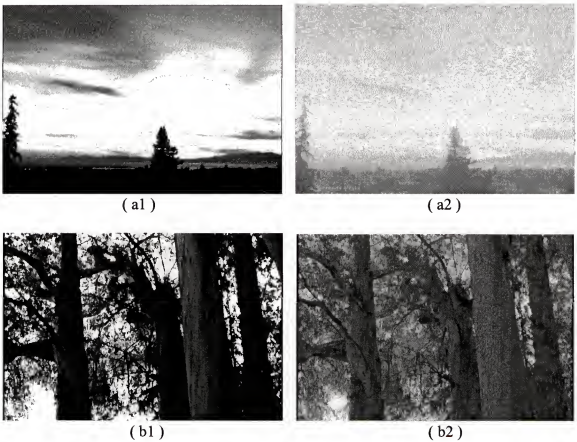
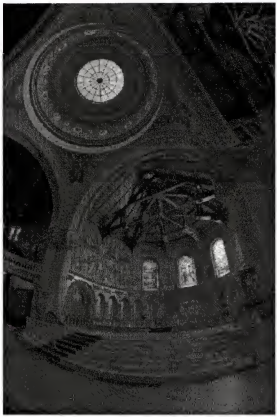



Figure 3.12 The left column shows the original wide dynamic range images: a1: 97dB, b1: 103dB, c1: 113dB, d1: 124dB, e1: 130dB, and f1: 168dB (source: www.debevec.org). The right column shows the corresponding sampled images with 8-bit logarithm luminance representation.

( c1 )

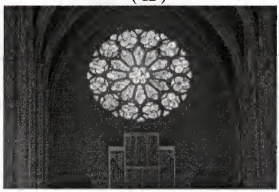


( c2 )

( d1 )

( d2 )

( e1 )



( e2 )

Figure 3.12 Continued.

( f1 )

( f2 )

Figure 3.12 Continued.

# CHAPTER 4
# SAMPLING STRATEGY

Optimal sampling strategies are discussed in this chapter. A sampling strategy consists of a sequence of reference voltages $V_{ref}$(k) and corresponding sampling times $T$(k). The combination of these two degrees of freedom will provide ultra-wide dynamic range imaging in video applications. Assume that there are a total of $N$ quantized levels of the photocurrent that are represented by $I(1)$, $I(2)$, ..., $I(N)$ where $I(1)$ and $I(N)$ are the largest and smallest photocurrents that are represented respectively. The corresponding sampling times are $T(1)$, $T(2)$, ..., $T(N)$ and reference voltages are $V_{ref}(1)$, $V_{ref}(2)$, ..., $V_{ref}(N)$. All pixels with the photocurrent value $I_{ph}$ that satisfies

$$I(k+1) \le I_{ph} < I(k) \qquad (k = 0,1,2,...N) \tag{4.1}$$

are assigned the integer ($k$+1). Here $I(0)$ is defined as $I(0) \equiv \infty$ and the corresponding $T(0)$ is defined as $T(0) \equiv 0$ to make equation (4.1) valid. $T(N)$ is a natural parameter to use since it corresponds to the maximal time to be allowed for the sampling which gives us the frame-rate. Here we also define

$$\Delta I(k) \equiv I(k+1) - I(k) \qquad (k = 1,2,..., N-1) \tag{4.2}$$

$$\Delta T(k) \equiv T(k+1) - T(k) \qquad (k = 1,2,..., N-1) \tag{4.3}$$

$$\Delta V_{ref}(k) \equiv V_{ref}(k+1) - V_{ref}(k) \qquad (k = 1,2,..., N-1) \tag{4.4}$$

and assume sufficient resolution of $V_{ref}$ can be obtained. Figure 4.1 shows the parameters and their relation.

Firstly, we discuss the constraint due to optimal signal-to-noise ratio (SNR) considerations. Following this, the application of linear quantization of light intensity and Weber's law are presented. In the discussion, we first explore the simple case by assuming there is no limitation on sampling time, and then we discuss the more important case with limited sampling time. Lastly, the Weber's law application in our design is presented.

Figure 4.1 The parameters and their relation.

### 4.1 Criterion Due To The Optimal SNR Consideration

The photogenerated charge collection rate (see chapter 2 for more details) is described by

$$I_{ph} = P_0 q\left(e^{-\alpha x1} - e^{-\alpha x2}\right)A \tag{4.5}$$

where $P_0$ is the expected value of incident photon flux (photons/cm$^3$·s) at the photodiode surface, the value the sensor is deigned to quantify, $q$ is electron charge, $A$ is the photodiode area and $x_1$ and $x_2$ denote the boundaries of the depletion layer (see Figure 2.2). Since the photocurrent $I_{ph}$ is proportional to $P_0$, the measurement of $I_{ph}$ indirectly quantifies the value of $P_0$ assuming the other parameters are constant. The arrival of photons can be described using a simple stream model. Its statistical property is described by Poisson distribution:

$$\Phi_n(t) = \frac{(P_0 t)^n}{n!} e^{-P_0 t} \tag{4.6}$$

where $\Phi_n(t)$ denotes the probability that $n$ photons arrive at the surface of the photodiode within time of $t$. In the measurement we expect $P_0 t$ photons to arrive at the surface of the photodiode within time $t$ since this is the expected value of the Poisson distribution.

Clearly the optimal sampling strategy will depend on the optimality criteria. One optimality criterion for digital imaging application is to maximize the SNR at each pixel. For a given technology and reset voltage, the total well charge of the reverse biased photodiode is

$$Q_{total} \equiv C_{ph} V_{res} \tag{4.7}$$

The noise analysis in Chapter 6 shows that the longer the integration is (before pixel saturation), the higher the SNR that can be achieved. Thus, a long integration time is desirable. For imaging applications, this implies the use of the whole well charge ($Q_{total}$) and therefore suggests the reference voltages should be set to zero for all time.

$$V_{ref,opt} = 0 \tag{4.8}$$

Note that this reference voltage may not be able to be implemented due to the limitation of input range and offset of comparators. A minimum reference voltage can be set as the lower boundary. Here we assume that reference voltage can reach 0 to simplify the analysis. More generally, we can replace 0 by a minimum reference voltage. Using the maximum well capacity implies that the corresponding firing time for the particular light intensity is

$$T_{opt} \equiv \frac{C_{ph}\left(V_{res} - V_{ref,opt}\right)}{I_{ph}} = \frac{C_{ph}V_{res}}{I_{ph}} \equiv \frac{\Psi}{I_{ph}} \tag{4.9}$$

The Full Well Charge (FWC) sampling constraint consists of equations (4.8) and (4.9).

Another constraint is the time domain minimum contrast constraint that is described by equation (4.10), where $\Delta T_{\min}$ is the worst case sampling interval and decided by hardware design.

$$\Delta T(k) \geq \Delta T_{\min} \qquad (k = 1,2,\ldots,N) \tag{4.10}$$

Giving these constraints, the major factors that affect sampling strategies in a time-based imaging system are

- reset voltage: $V_{res}$
- capacitance at the cathode of photodiode: $C_{ph}$
- maximum possible sampling time: $T_{max}$
- worst case sampling interval (worst case readout time per frame): $\Delta T_{\min}$

To explore the effect of these factors, sampling strategies with and without $T_{max}$ limitation are presented in the following sections. Assume the reference voltage of the first sample is set to 0 so that the first sampling point is decided by $\frac{\Psi}{I(1)}$ and become a known value. The sequence of reference voltages and sampling times are determined respectively.

### 4.1.1 Sampling Strategy Without Sampling Time Limitation

Extending the FWC to the whole imaging procedure, a sampling strategy where the reference voltages are zeros for all the samplings ($V_{ref}(k) = V_{ref}(k)_{opt} = 0$) and the

sampling time is set to the corresponding optimal value $T(k) = T(k)_{opt}$ can be obtained. This is shown in Figure 4.2.

Since a constant reference voltage is adopted in the sampling strategy, the characteristics of the sampling are represented by

$$I(k)T(k) = \Psi \tag{4.11}$$

where $\Psi$ denotes the constant value of $C_{ph}V_{res}$. The sampling interval

$$\Delta T(k) = T(k+1) - T(k) \qquad (k = 1,2,...N) \tag{4.12}$$

$$= \Psi \frac{I(k) - I(k+1)}{I(k)I(k+1)} \tag{4.13}$$

$$\equiv \frac{\Psi \Delta I(k)}{I(k)I(k+1)} \tag{4.14}$$

must satisfy

$$\Delta T(k) \geq \Delta T_{\min} \qquad (k = 1,2,..., N)$$

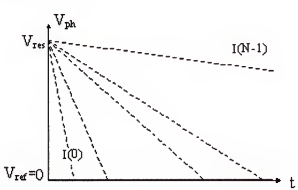



Figure 4.2 FWC constraint requires a constant reference voltage of 0.

For the cases where some $\Delta T(k)$'s are smaller than $\Delta T_{\min}$, $\Delta T(k)$ has to be increased until the time domain minimum contrast law is satisfied. This results in a partial use of the whole well charge and thus, an increase of the reference voltage. The sampling strategy is completed with the sacrificing of well charge and SNR. Generally, in

the high light intensity range, the denominator of Equation (4.14) is much larger than the numerator and thus $\Delta T(k)$ is small. As light intensity decreases, $\Delta T(k)$ becomes larger. So $\Delta T(k) < \Delta T_{\min}$ usually occurs in the high light intensity range. Here, we assume this is always true for the applications of different quantization. Assume

$$\Delta T(k) < \Delta T_{\min}, \quad \Delta T(k+1) \geq \Delta T_{\min}$$

Then

$$I(k)T'(k)$$

$$\equiv I(k)(T(k+1) - \Delta T_{\min}) \tag{4.15}$$

$$= C_{ph}(V_{res} - V_{ref}(k)) < \Psi \tag{4.16}$$

where $T'(k) = T(k+1) - \Delta T_{\min}$. To use the total well charge as much as possible, the $V_{ref}(k)$ in (4.16) should be as small as possible. Thus, $\Delta T(k)$ should be increased to just $\Delta T_{\min}$ as shown in (4.15). This rule should be applied to all the $\Delta T(k)$'s that are smaller than $\Delta T_{\min}$ in the implementation of sampling strategy. Figure 4.3 shows the possible reference voltage varying in the high light intensity range.

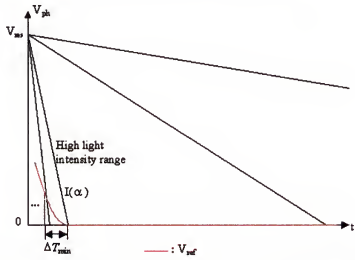



Figure 4.3 Possible reference voltages varying in the high light intensity range

Here we name the first sampling point at which $V_{ref}(k) = 0$ as $\alpha$ (see Figure4.3). Clearly if there are no other constraints, the optimal sampling strategy would make $V_{ref}$ always equal to zero. With timing constraints, SNR optimality gets more complicated however it can be argued that the resulting sampling strategy cannot be improved upon. If we attempt to improve the SNR at any light intensity by decreasing $V_{ref}$ at that point then a constraint will be violated (either the time domain minimum contrast constraint or $V_{ref}$ would go below zero). The strategy, therefore, it still optimal given the constraints. So the sampling strategy without sampling time limitation can be summarized as

- For $1 \le k \le \alpha$

$$\begin{cases} T(k) = T(\alpha) - (\alpha - k)\Delta T_{\min} \\ V_{ref}(k) = V_{res} - \dfrac{I(k)T(k)}{C_{ph}} \end{cases} \tag{4.17}$$

- For $\alpha < k \le N$

$$\begin{cases} T(k) = T(k)_{\max} = \dfrac{\Psi}{I(k)} \\ V_{ref}(k) = 0 \end{cases} \tag{4.18}$$

Notice that if $T$(k) in equation (4.17) is less than 0, it becomes invalid. This will result in less resolution. A dedicate design should avoid this problem.

**4.1.2 Sampling Strategy With A Limited Sampling Time**

In most application cases, a finite sampling time is required as shown in Figure 4.4. For example, in video applications the maximum sampling time is 33ms in order to provide a frame rate of 30fps.

A simple option is to still use a constant reference voltage as we have done in last section. However, if we still choose zero as the reference voltage, those photocurrents

having a magnitude less than $I_{dia}$ (see Figure 4.4) will never trigger the comparator. At $T_{\max}$, the analog voltages could be read out but this requires ADC and all of the additional analog processing as in conventional CMOS imagers. This solution is not preferred since the advantages of time-based CMOS imagers cannot be totally realized and this mixed-signal processing results in a larger pixel size (two sets of circuits) and much complicated sampling and reconstruction. If no analog readout is implemented, all the $I_{ph}$'s less than $I_{dia}$ will be assigned the value of $I_{dia}$ and thus, the dynamic range is dramatically reduced approximately halved in dB). There is no advantage of ultra-wide dynamic range imaging in this case. Another idea is to enable every pixel to trigger in the limited time by setting the reference voltage to be a constant voltage slightly below $V_{res}$. However, this approach has two major problems. First, well charge is badly used and second the sampling interval for high light intensity is much smaller than $\Delta T_{\min}$ and therefore it is infeasible. This can be illustrated by

$$\begin{cases} I_{ph}(k)T(k) = C_{ph}(V_{res} - V_{ref}) \\ I_{ph}(k+1)T(k+1) = C_{ph}(V_{res} - V_{ref}) \end{cases}$$

$$\Rightarrow \quad \Delta T(k) = T(k+1) - T(k) = \frac{C_{ph}(V_{res} - V_{ref})\left(I_{ph}(k) - I_{ph}(k+1)\right)}{I_{ph}(k)I_{ph}(k+1)}$$

$$\Rightarrow \quad \Delta T(k) \approx \frac{C_{ph}V_{off}\Delta I_{ph}(k)}{I_{ph}(k)I_{ph}(k+1)}$$

$$= \frac{I_{ph,\min}T_{\max}\Delta I_{ph}(k)}{I_{ph}(k)I_{ph}(k+1)}$$

$$\approx \frac{10^{-15}10^{-3}\Delta I_{ph}(k)}{10^{-9}10^{-9}} \qquad \text{for high light intensity range}$$

$$< 10^{-9}\,\text{sec} << \Delta T_{\min} \approx order \quad of \quad \mu s$$

A strategy with constant reference voltage of value between 0 and ($V_{res}$ - $V_{off}$) will suffer reduced dynamic range and bad time domain contrast. Thus, a varying reference voltage is preferred over any constant reference voltage strategy. The general direction of reference voltage change is from the bottom left corner to the top right corner as shown in Figure 4.4.

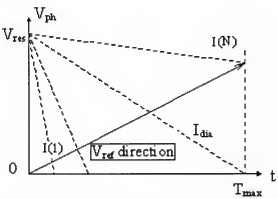



Figure 4.4 Relation between the photodiode voltage and the integration time for different photocurrents in the application with finite sampling time. The general direction of reference voltage change for imaging with finite sampling time is from left bottom to right top.

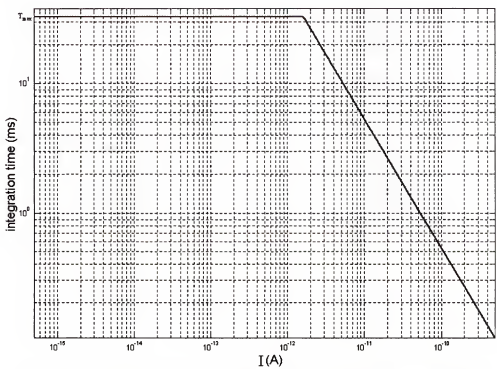



Figure 4.5 The upper boundary of sampling time for video applications.

It is clear that all the photocurrents having a magnitude less than $I_{dia}$ have maximum sampling time of $T_{\max}$. While all the photocurrents having a magnitude greater than $I_{dia}$ have maximum sampling time less than $T_{\max}$. So the upper boundary of sampling time can be decided by

$$\begin{cases} T(k)_{\max} = \dfrac{C_{ph}V_{res}}{I(k)} & for \quad I(k) \geq I_{dia} \\ T(k)_{\max} = T_{\max} & for \quad I(k) < I_{dia} \end{cases} \tag{4.19}$$

The upper boundary of sampling time $T(k)_{\max}$ for video application is shown in Figure 4.5 where $C_{ph}$ = 14.8fF, $V_{res}$ = 3V, $I_{ph,\max}$ = 0.5nA and $I_{ph,\min}$ = 0.5fA and $I_{dia}$ can be found to be 1.62pA.

In order to maximally use the total well charge, we should use the $T(k)_{\max}$ for each $I(k)$. However, $T_{\max}$ is shared by the $I(k) \leq I_{dia}$ and only one of them can use the maximum sampling time $T_{\max}$. To solve this problem, $T_{\max}$ is assigned to $T(N)$ and others should satisfy

$$\begin{cases} T(N) = T_{\max} \\ T(k) = T_{\max} - (N-k)\Delta T_{\min} \quad (k \leq N-1) \end{cases} \tag{4.20}$$

Thus, the reference voltages will not be zero for the samplings where $I(k) \leq I_{dia}$. The varying of reference voltages in this range can be described by

$$\begin{aligned} V_{ref}(k) &= V_{res} - \frac{I(k)T(k)}{C_{ph}} \\ &= V_{res} - \frac{I(k)}{C_{ph}}\left[T_{\max} - (N-k)\Delta T_{\min}\right] \end{aligned} \tag{4.21}$$

As $k$ changes from 1 to $N$, equation (4.21) indicates that the reference voltages monotonically increase. Here we will define the point where $V_{ref}(k) = 0$ in equation

(4.21) as $\beta$. Obviously, $I(\beta) > I_{dia}$. Figure 4.6 shows a possible reference voltage varying in the low light intensity range as well as the possible reference voltage varying in the high light intensity range (see Figure 4.3).

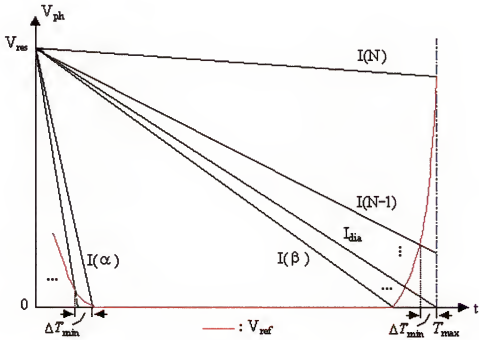



Figure 4.6 Possible reference voltage varying in the low light intensity range as well as the possible reference voltage varying in the high light intensity range.

The sampling strategy can be summarized in the following equations.

- For $\beta \le k \le N$,

$$\begin{cases} T(k) = T_{\max} - (N-k)\Delta T_{\min} \\ V_{ref}(k) = V_{res} - \dfrac{I(k)T(k)}{C_{ph}} \end{cases} \tag{4.22}$$

- For $\alpha < k < \beta$,

$$\begin{cases} T(k) = T(k)_{\max} = \dfrac{\Psi}{I(k)} \\ V_{ref}(k) = 0 \end{cases} \tag{4.23}$$

- For $1 \le k \le \alpha$

$$\begin{cases} T(k) = T(\alpha) - (\alpha - k + 1)\Delta T_{\min} \\ V_{ref}(k) = V_{res} - \dfrac{I(k)T(k)}{C_{ph}} \end{cases} \tag{4.24}$$

For different sampling applications, we need to find the $\alpha$ and $\beta$ then the sampling strategy can be obtained following equations (4.22)-(4.24). A pseudo-code is presented below to illustrate the processing.

Next, we will explore the application of linear sampling and the application of Weber's law respectively to see how well the general sampling strategy we discussed here can be implemented to particular applications.

```
For k=1:N
        If ΔT(k) > ΔT_min                                  % find α
                α = k ;
                break;
        end
end

From equation (4.21), let V_res − I(β)/C_ph [T_max − (N − β)ΔT_min] = 0
and find β.

Based on equation (4.22) – (4.24), decide the reference voltages and
sampling times for the sampling strategy.
```

## 4.2 Sampling Strategy For The Applications Of Linear Quantization Of Light Intensity

The most straightforward application of sampling is to uniformly quantize wide dynamic range photocurrent. The photocurrents selected under this application satisfy

$$I(k) = I(1) - (k-1)d \qquad (k = 1,2,...N) \tag{4.25}$$

where $d$ denotes the brightness step and can be obviously described by

$$d = \frac{I(1) - I(N)}{N-1} \approx \frac{I(1)}{N-1} \tag{4.26}$$

First of all we will discuss the application without sampling time limitation. Then we will discuss the cases with limited sampling time. Sampling strategies for both cases will be derived and summarized.

### 4.2.1 Sampling Strategy Without Sampling Time Limitation

The sampling strategy of FWC indicates that the relation between $I$ and $t$ is described by equation (4.7) and the $I$-$t$ curve is illustrated in Figure 4.7.

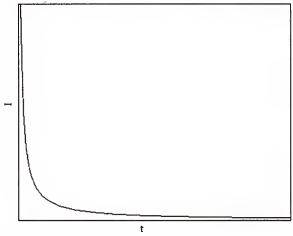



Figure 4.7 The relation of $I$ and $t$ described by equation (4.7).

Thus, from equation (4.13) we can obtain

$$\Delta T(k) = \Psi \frac{d}{I(k)I(k+1)} \qquad (k = 1,2,\ldots,N) \tag{4.27}$$

$$= \Psi \frac{d}{[I(1) - (k-1)d][I(1) - kd]} \tag{4.28}$$

Substituting (4.26) we can obtain

$$\Delta T(k) = \frac{\Psi / d}{k^2 - (2N-1)k + N(N-1)} \tag{4.29}$$

The time domain minimum contrast law (equation (4.14)) requires

$$\Delta T(k) = \frac{\Psi / d}{k^2 - (2N-1)k + N(N-1)} \geq \Delta T_{\min} \tag{4.30}$$

This is a quadratic equation with the maximum value of $\Delta T(k)$ at

$$k = N - \frac{1}{2} \tag{4.31}$$

and thus, $\Delta T(k)$ is monotonically increasing with $k \in [1, N]$. Let $\Delta T(k) = \Delta T_{\min}$, we can solve the equation of (4.30) and obtain

$$\alpha = N - \sqrt{\frac{\Psi}{\Delta T_{\min} d}} \tag{4.32}$$

The reason of the existence of α is that the slope of *I-t* curve at high light intensity range is too sharp to provide the required minimum sampling interval. The slope can be found to be

$$S = \frac{dI}{dt} = -\frac{\Psi}{t^2} = -\frac{I^2}{\Psi} \tag{4.33}$$

It is proportional to $I^2$, which is very large in high light intensity range. To satisfy the time domain minimum contrast law, we need to reduce the slope in high light intensity range. The sampling strategy for the samples between 0 and α should satisfy

$$T(k) = T(\alpha) - (\alpha - k)\Delta T_{\min} \qquad (k = 1, 2, \ldots, \alpha) \tag{4.34}$$

Thus, the relation of *I* and *t* for samples between 1 and α can be described by

$$I(k) = \left[ I(1) - d\left( \alpha - 1 - \frac{T(\alpha)}{\Delta T_{\min}} \right) \right] - \frac{d}{\Delta T_{\min}} T(k) \tag{4.35}$$

The slope is now $S = -\frac{d}{\Delta T_{\min}}$, whose absolute value is smaller that that of equation (4.33). Let

$$\frac{\left[I(1)-(k-1)d\right]^2}{\Psi}=\frac{d}{\Delta T_{\min}} \tag{4.36}$$

We can obtain

$$k=N-\sqrt{\frac{\Psi}{d\Delta T_{\min}}} \tag{4.37}$$

which is exactly the α. Thus, the revised straight line should be the tangent of the hyperbola at point α. This is displayed in Figure 4.9(a). When α is small, $I(1)$ can still be covered by the revised straight line and the resolution can be kept as shown in Figure 4.9(a). However, if α is large enough, some of the $I$(k)'s no longer can be covered and resolution is lost at these $I$(k)'s as shown in Figure 4.8(b).

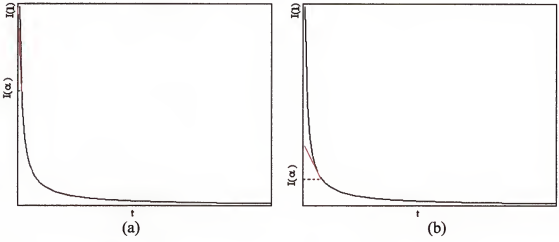



Figure 4.8 The sampling strategy in the high light intensity range. The slope change (red straight line) in the high light intensity range as described by equation (4.35). (a) When α is small, $I(1)$ can be covered. (b) When α is large, some of the $I$(k)'s no longer can be covered and resolution is lost at these $I$(k)'s.

The reference voltage satisfy

$$\begin{cases} C_{ph}\left(V_{res}-V_{ref}(k)\right)=I(k)T(k) & (k=1,2,\ldots,\alpha) \\ V_{ref}(k)=0 & (k=\alpha+1,\alpha+2,\ldots,N) \end{cases} \tag{4.38}$$

Let $A_1 \equiv I(1)+d$, $A_2 \equiv T(\alpha)-\alpha\Delta T_{\min}$, we can obtain

$$V_{ref}(k) = \frac{1}{C_{ph}}\left[\Psi - A_1 A_2 - k\left(A_1 \Delta T_{\min} - A_2 d\right) + k^2 d \Delta T_{\min}\right] \tag{4.39}$$

Combining the above results, we can obtain the sampling strategy for the application of linear sampling. This is concluded in the following equations.

- For sampling points $1 \le k \le \alpha$,

$$\begin{cases} T(k) = T(\alpha) + (\alpha - k)\Delta T_{\min} \\ V_{ref}(k) = \dfrac{1}{C_{ph}}\left[\Psi - A_1 A_2 - k\left(A_1 \Delta T_{\min} - A_2 d\right) + k^2 d \Delta T_{\min}\right] \end{cases} \tag{4.40}$$

- For sampling points $\alpha < k \le N$

$$\begin{cases} T(k) = \dfrac{\Psi}{I(k)} = \dfrac{\Psi}{I(1) + (k-1)d} \\ V_{ref}(k) = 0 \end{cases} \tag{4.41}$$

Figure 4.9(a) shows an example of $V_{ref}$ varying with the parameters of $C_{ph}$ = 17.8fF, $V_{res}$ = 3V, $I_{ph,\max}$ = 0.5nA, $I_{ph,\min}$ = 0.5fA, $N$=256 and $\Delta T_{\min}$ =1.5μs.

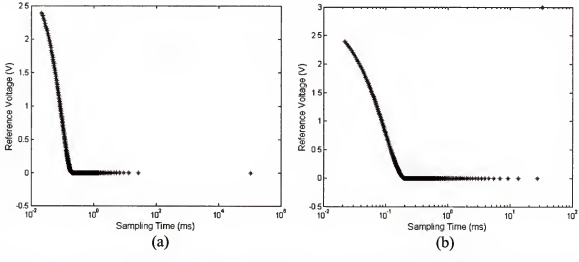



Figure 4.9 A sampling strategy example: reference voltage varying with sampling time for the application of linear quantization of light intensity. (a) Without sampling time limitation. (b) The maximum sampling time is 33ms.

### 4.2.2 Time-Based Sampling With A Limited Sampling Time

In this section, we need to find the point of $\beta$. As we have discussed in section 4.1, we let $V_{ref}(k)$ in equation (4.21) equal to 0 and can obtain

$$V_{res} - \frac{I(\beta)}{C_{ph}}\left[T_{\max} - (N-\beta)\Delta T_{\min}\right] = 0 \tag{4.42}$$

where $I(\beta) = I(1) - (\beta - 1)d$. Let $A_3 \equiv T_{\max} - N\Delta T_{\min}$, we can obtain

$$\beta = k_{\min,\beta} + \frac{\sqrt{\left(A_1\Delta T_{\min} + A_3 d\right)^2 - 4d\Delta T_{\min}\Psi}}{2d\Delta T_{\min}} \tag{4.43}$$

where $k_{\min,\beta} = \dfrac{A_1\Delta T_{\min} - A_3 d}{2d\Delta T_{\min}}$. Notice that

$$N \le \frac{T_{\max}}{\Delta T_{\min}} + 1 \tag{4.44}$$

If $T_{\max} = 33ms$ and $\Delta T_{\min} = 5\mu s$, we can at most have 6601 sampling points and thus, at most 13 bits can be used in the representation of the sampling.

The reference voltage at $\beta \le k \le N$ can be calculated from

$$V_{ref}(k) = V_{res} - \frac{\left(I(1) - (k-1)d\right)T(k)}{C_{ph}} \tag{4.45}$$

where $T(k) = T_{\max} - (N-k)\Delta T_{\min}$. We can obtain

$$V_{ref}(k) = \frac{1}{C_{ph}}\left[\Psi - A_1 A_3 - k\left(A_1\Delta T_{\min} - A_3 d\right) + k^2 d\Delta T_{\min}\right] \tag{4.46}$$

Compare equation (4.46) with equation (4.39), we can see that they have same expression except that $A_3$ replaces $A_2$. In summary, combined with the results in section 4.2.1, the sampling strategy is described by the following equations.

- For sampling points $1 \le k \le \alpha$,

$$\begin{cases} T(k) = T(\alpha) - (\alpha - k)\Delta T_{\min} \\ V_{ref}(k) = \dfrac{1}{C_{ph}}\left[\Psi - A_1 A_2 - k\left(A_1 \Delta T_{\min} - A_2 d\right) + k^2 d \Delta T_{\min}\right] \end{cases} \tag{4.40}$$

- For sampling points $\alpha < k < \beta$

$$\begin{cases} T(k) = \dfrac{\Psi}{I(k)} = \dfrac{\Psi}{I(1) - (k-1)d} \\ V_{ref}(k) = 0 \end{cases} \tag{4.41}$$

- For $\beta \le k \le N$,

$$\begin{cases} T(k) = T_{\max} - (N - k)\Delta T_{\min} \\ V_{ref}(k) = \dfrac{1}{C_{ph}}\left[\Psi - A_1 A_3 - k\left(A_1 \Delta T_{\min} - A_3 d\right) + k^2 d \Delta T_{\min}\right] \end{cases} \tag{4.47}$$

Using the same parameters of Figure 4.9(a), an example with maximum sampling time of 33ms is shown in Figure 4.9(b).

**4.2.3 Summary And Discussion**

When there is no sampling time limitation, the sampling strategy of the linear quantization application needs to define $\alpha$ first (equation (4.32)). When $\Delta T_{\min}$ becomes larger, $\alpha$ becomes larger. This will cause loss of resolution and thus, a small $\Delta T_{\min}$ is required in this application. For example, 1.5us is used in the example shown in Figure 4.9. To check whether the resolution may be lost, a tangent line can be drawn at point $\alpha$. If this tangent line can cover $I(1)$, then resolution will not be lost. Otherwise, the resolution will be lost. For the application with finite sampling time, the sampling strategy is described by the equations (4.40), (4.41) and (4.47). In the next section, we will explore a nonlinear quantization application that is quite different from the linear quantization application. We will see that although the results are different, the exact same method is implemented.

## 4.3 Sampling Strategy For The Application Of Weber's Law

As discussed in Chapter 3, Weber's law corresponds well with human visual response, whose contrast threshold is constant percentage over a wide range of adaptation luminance. The photocurrents selected under the Weber's law satisfy

$$I(k) = r^{k-1} I(1) \tag{4.48}$$

where $k$=1, 2, 3, …, $N$ and $N$ is the total number of sampling points. $r$ denotes the relative percentage brightness step and is a value less than 1 and generally greater than 0.95. For Weber's law, the logarithm domain quantization error, $\log I(k) - \log I(k-1)$, is a constant, which is equal to $\log r$.

### 4.3.1 Sampling Strategy Without Sampling Time Limitation

From (4.11) and (4.48), we can obtain

$$r^{k-1} \frac{T(k)}{T(1)} = 1 \quad (k = 1,2,\ldots,N) \tag{4.49}$$

This implies

$$T(k) = \frac{T(1)}{r^{k-1}} \quad (k = 1,2,\ldots,N) \tag{4.50}$$

Equation (4.50) shows that the integration time for $I_{ph}(k)$ is exponentially increasing. Further more,

$$\begin{aligned} \Delta T(k) &= T(k+1) - T(k) \\ &= T(1)\left(\frac{1}{r^k} - \frac{1}{r^{k-1}}\right) \\ &= \frac{T(1)}{r^{k-1}}\left(\frac{1}{r} - 1\right) \\ &= T(k)\left(\frac{1-r}{r}\right) \quad (k = 1,2,\ldots,N-1) \end{aligned} \tag{4.51}$$

If no reference voltage adjustment can be involved, the shortest sampling interval derived from equation (4.51) must satisfy

$$\Delta T(1) = T(1)\left(\frac{1-r}{r}\right) \geq \Delta T_{\min} \tag{4.52}$$

$$\Rightarrow \quad \frac{1-r}{r} \geq \frac{\Delta T_{\min}}{T(1)} \tag{4.53}$$

$$\Rightarrow \quad r \leq \frac{1}{1+\frac{\Delta T_{\min}}{T(1)}} \tag{4.54}$$

Generally, $r \geq 0.95$, so

$$0.95 \leq r \leq \frac{1}{1+\frac{\Delta T_{\min}}{T(1)}}$$

$$\Rightarrow \quad T(1) \geq \frac{r}{1-r}\Delta T_{\min} \geq 19\Delta T_{\min} \tag{4.55}$$

If $\Delta T_{\min}$ is chosen to be 5 μs, then

$$T(1) \geq 95\mu s$$

For the cases where (4.55) is violated, $\alpha$ exists. Let

$$T(\alpha) = \frac{\Psi}{I(1)r^{\alpha-1}} = \frac{r}{1-r}\Delta T_{\min} \tag{4.56}$$

we can obtain

$$(\alpha-1)\ln r = \ln\frac{(1-r)\Psi}{rI(1)\Delta T_{\min}} \tag{4.57}$$

So

$$\alpha = \frac{\ln\frac{(1-r)\Psi}{I(1)\Delta T_{\min}}}{\ln r} \tag{4.58}$$

The operation of sensing within [$T(1)$, $T(\alpha)$] is described by

$$I(k)T(k) = C_{ph}\left(V_{res} - V_{ref}(k)\right) \quad (1 \le k \le \alpha) \tag{4.59}$$

From equation (4.59), we can obtain

$$r^{k-\alpha}\frac{T(k)}{T(\alpha)} = \frac{\left(V_{res} - V_{ref}(k)\right)}{V_{res}} \quad \left(1 \le k \le \alpha\right) \tag{4.60}$$

$$= \left(1 - \frac{(1-r)(\alpha-k)}{r}\right) r^{k-\alpha} \tag{4.61}$$

So

$$V_{ref}(k) = V_{res}\left(1 - r^{k-\alpha} + (1-r)(\alpha-k)r^{k-\alpha-1}\right) \tag{4.62}$$

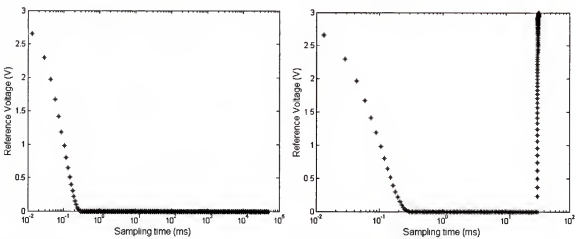



Figure 4.10 The varying reference voltages in the application of Weber's law. Notice the bending-up of the reference voltage in high and low light intensity range. (a) Without sampling time limitation. (b) The maximum sampling time is 33ms.

The sampling strategy is summarized as

- For sampling points $1 \le k \le \alpha$,

$$\begin{cases} T(k) = T(\alpha) - (\alpha - k)\Delta T_{\min} \\ V_{ref}(k) = V_{res}\left(1 - r^{k-\alpha} + (1-r)(\alpha-k)r^{k-\alpha-1}\right) \end{cases} \tag{4.63}$$

- For sampling points $\alpha < k \le N$

$$\begin{cases} T(k) = \dfrac{\Psi}{I(k)} = \dfrac{\Psi}{I(1) + (k-1)d} \\ V_{ref}(k) = 0 \end{cases} \tag{4.64}$$

Once more, note that the increased reference voltages results in the scarificing of total well charge. Figure 4.10(a) shows an example with the existence of $\alpha$, where $C_{ph}$ = 17.8fF, $V_{res}$ = 3V, $I_{ph,\max}$ = 0.5nA, $I_{ph,\min}$ = 0.5fA, $N$=256 and $\Delta T_{\min}$ =15μs.

### 4.3.2 Sampling Strategy With A Limited Sampling Time

Similarly, we need to find the value of $\beta$. Let $V_{ref}(k)$ in equation (4.21) equal to 0, we can obtain

$$V_{res} = \frac{1}{C_{ph}} I(1) r^{\beta-1} \left[ T_{\max} - (N-\beta)\Delta T_{\min} \right] \tag{4.65}$$

$$\Rightarrow \quad T(\beta) = \frac{C_{ph} V_{res}}{I(1) r^{\beta-1}} = T_{\max} - (N-\beta)\Delta T_{\min} \tag{4.66}$$

It is impossible to derive a close form solution for $\beta$. Here we will find the intersections of the two functions

$$y(\beta) = \frac{\Psi}{I(1)} r^{-\beta+1} \tag{4.67}$$

$$y(\beta) = T_{\max} - N\Delta T_{\min} + \beta\Delta T_{\min} \tag{4.68}$$

Graphically we can find the solution of $\beta$. Figure 4.11 shows the results with $N$ = 256, $r$ = 0.95. When $\Delta T_{\min}$ is small, there exist intersections of the straight line (equation (4.68)) and the exponential curve (equation (4.67)) and therefore $\beta$ exists (if there exist two intersection points, the larger one is selected). As the $\Delta T_{\min}$ increases, the slope of the straight line increases too. After the tangency with the exponential curve, equation (4.68) o longer intersects with equation (4.67).

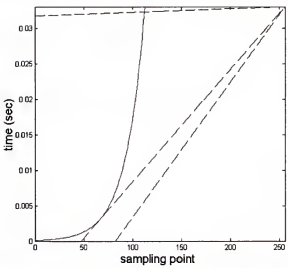



Figure 4.11 Intersection of equation (4.67) (solid line) and (4.68) (dash line).

If $\beta$ exists, then the sampling strategy is described by

- For sampling points $1 \le k \le \alpha$,

$$\begin{cases} T(k) = T(\alpha) - (\alpha - k)\Delta T_{\min} \\ V_{ref}(k) = V_{res}\left(1 - r^{k-\alpha} + (1-r)(\alpha - k)r^{k-\alpha-1}\right) \end{cases} \tag{4.69}$$

- For sampling points $\alpha < k < \beta$

$$\begin{cases} T(k) = \dfrac{\Psi}{I(k)} = \dfrac{\Psi}{I(1) + (k-1)d} \\ V_{ref}(k) = 0 \end{cases} \tag{4.70}$$

- For sampling points $\beta \le k \le N$

$$\begin{cases} T(k) = T_{\max} - (N - k)\Delta T_{\min} \\ V_{ref}(k) = V_{res} - \dfrac{1}{C_{ph}} I_{ph}(1) r^{k-1}\left[T_{\max} - (N-k)\Delta T_{\min}\right] \end{cases} \tag{4.71}$$

An example is shown in Figure 4.10(b) with the maximum sampling time of 33ms. Notice that when $\Delta T_{\min}$ is too large, $\beta$ does not exist. This results in decrease of dynamic range. A very large array size may cause nonexistence of $\beta$. Then row parallel output or on-chip memory scheme should be implemented to solve this problem.

Similarly, the resolution is constrained by

$$N \le \frac{T_{\max}}{\Delta T_{\min}} + 1 \tag{4.72}$$

For video applications, if $N$ = 256, then $\Delta T_{\min} \approx 129\mu s$. This is the largest sampling interval we can use. Equation (4.72) also implies that if the design requires $\Delta T_{\min}$ = 5us and the maximum integration time is 33ms, then $N \le 6601$. This corresponds to at most 13 bits representation of the sampling points.

**4.3.3 Summary**

In a time-based sampling digital imaging system, $\Delta T_{\min}$ is decided by the design clock frequency and array size. The ratio $r$ for Weber's law application generally has a value greater than 0.95. These two parameters require that under FWC

$$T(1) \ge \frac{r}{1-r}\Delta T_{\min}. \tag{4.73}$$

Thus, the maximum detectable photocurrent under FWC turns out to be

$$I(1) = \frac{(1-r)C_{ph}V_{res}}{r\Delta T_{\min}}. \tag{4.74}$$

Since $C_{ph}$, $V_{res}$ and $\Delta T_{\min}$ are decided as long as the system is built up, $r$ is the variable we only have flexibility on. By selecting a proper $r$, we can avoid adjusting reference voltage in the high light intensity range.

If $\alpha$ exists, the minimum time domain contrast constraint should be involved in the high light intensity range and the reference voltage has to bend up towards the high light intensity direction as described by equation (4.63). Similarly, if $\beta$ exists, the minimum time domain contrast constraint should be involved in the low light intensity

range and the reference voltage has to bend up towards the low light intensity direction as decide by equation (4.71).

The $\alpha$ and $\beta$ in the Weber's law application is general less than that in the linear quantization application. This is because that the Weber's law application provides more resolution in low light intensity range and roast resolution in high light intensity range, while the linear quantization application works in an opposite way.

### 4.4 Discussion

Based on the two constraints, the time domain minimum contrast constraint and full well charge capacity constraint, we can develop optimal sampling strategies for maximizing SNR for different applications. The major task is to find the points of $\alpha$ and $\beta$. They vary for different applications. Generally, they should not overlap each other. For the cases where $\alpha$ and $\beta$ overlap each other, the time-varying reference voltage is chosen from the higher value of the possible reference voltages. Using equations (4.22)-(4.24), we can obtain the sampling time sequence and corresponding reference voltages. This method can also be extended to find the linear mapping of sampling data to the original image as illustrated by Astrom et al. [53]. The linear mapping application in near-sensor image processing is one particular case of the optimal sampling strategy.

Since multiple sampling is required in the time-based sampling technique, multiple frames need to be read off chip. In next chapter, we will discuss the readout schemes for efficiently realizing the time-based sampling.

# CHAPTER 5
# READOUT SCHEMES

## 5.1 Challenges And Methods

At each sampling time instance of the time-based CMOS imagers, we face the problem of reading out the information from the whole frame and therefore, the readout scheme must be designed to be fast enough. In other words, the readout speed is the crucial issue for the achievable dynamic range. The conventional high-end (the power supply voltage, see Chapter 3) constraint of dynamic range is now replaced by the readout speed.

We have mentioned that DPS pixels should be involved in the time-based CMOS imager design. When $V_{ph}$ drops under the reference voltage $V_{ref}$, the comparator inside the DPS pixel toggles. When the comparator toggles, we define the pixel as a "fired" pixel. Otherwise, we define the pixel as an "unfired" pixel. All the "fired" pixels at a sampling time instance will be assigned the same gray scale level. One of the simplest readout schemes is to read out the whole pixel array at each sampling time instance, that is, read out the whole frame. For example, in the literature [20] the whole frame is scanned and read out 256 times and thus, each pixel is represented by 256 bits. There exist three disadvantages in this simple readout scheme. First, a very high frequency clock is required to scan the whole array within a short time interval. For example, to read out a QCIF (176×144 pixels) image array within 5us requires clock frequencies as fast as 5GHz if no parallel processing is adopted. When the array size increases, the

required frequency becomes even higher, which is clearly infeasible. Second, since each pixel is read off the chip 256 times, the power consumption of the architecture is much more than what it can be if each pixel was read off the chip just once.

Different readout schemes can be implemented for different size pixel arrays. In general, we can divide the different readout schemes to two categories: on-chip memory readout schemes and off-chip memory readout schemes. Here we will discuss the off-chip memory readout schemes first. Generally, let's use $f_{max}$ to denote the highest available frequency. Here, $f_{max}$ = 66MHz is used as an example. Without losing generality, assume a square pixel array (size of N×N) as our objective.

The worst case of the readout process in a time-based CMOS imager is that the whole array should be read out within the shortest time interval (e.g., 5us). This limits the maximal size of a pixel array to be 18×18 pixels if no parallel processing is involved. Thus, to read out the whole frame greater than 18×18 pixels off chip in one sampling interval is not feasible. Fortunately, in general for typical images, one gray scale level in a 256 gray scale picture usually accounts for less than 2% of the total number of pixels as shown by VanRullen et. al. [57]. Considering this and if only those pixels that have the same gray scale level are read out, the maximal size of the pixel array can be as large as 128×128 pixels. So, the readout scheme that will only read out the information of those pixels that have "fired" at the sampling time instance has the advantage of improving the performance of time-based CMOS imagers. Figure 5.1 shows the block diagram of a CMOS imager with absolute address coding (this will be discussed in detail in section 5.2.1) readout scheme. The symbol inside the dash box represents digital pixel sensor (DPS) in which a comparator is typically included (see Chapter 6 for detail).

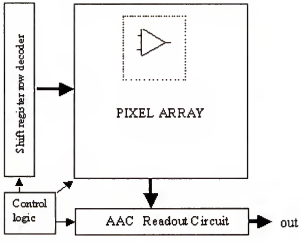



Figure 5.1 Block diagram of a CMOS imager with absolute address coding readout scheme.

If no parallel processing is involved, when the size of the pixel array increases and exceeds the limitation of $f_{max}$, an on-chip memory readout scheme can be implemented. Here, if real-time processing is not required (for example in the application of still image capture), but high space resolution imaging (for example 2048×2048 pixels) is required, an architecture with on-chip N×N×M-bit memory is suitable for realizing the ultra-wide dynamic range imaging, where each pixel value is represented by M bits. Here let's use M=8 bits as an example. The digital numbers 255, 254, 253, …, 0 represent the time instances $t_0$, $t_1$, $t_2$,…,$t_{255}$, respectively. Between two adjacent time instances, the rows of the pixel array are scanned one by one, as well as the corresponding rows in the on-chip memory array. The "fired" pixels in each row (assign logic value "1") enable the "write" processing of writing the digital numbers to the corresponding memory cells in the memory array. For instance, at time instance $t_0$, the "255" is written to the memory array. At time instance $t_1$, "254" is written to the memory array, and so on. After the whole image is captured, the memory data that represents the indices of the sampling time, can be read off the chip. Since only rows are scanned and

all the columns are processed in parallel, we do not need a fast clock. Figure 5.2 illustrates the block diagram of a CMOS imager with on-chip memory.

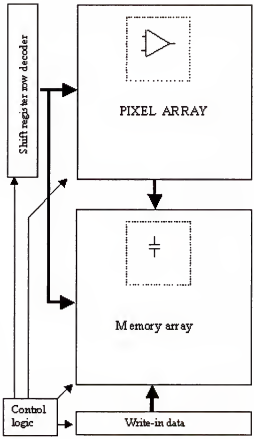



Figure 5.2 Block diagram of a CMOS imager with on-chip memory.

One of the advantages of these proposed schemes is that they avoid the high-speed clock requirement. For example, in the above description, 66MHz is used. In addition, each pixel is read off the chip only once and this results in the lowest power consumption at a given condition. With the proposed readout scheme in a time-based CMOS imager, an ultra-wide dynamic range and high space resolution imaging can also be achieved.

The obvious drawback for CMOS imagers with on-chip memory is that chip area increases. This is a classic tradeoff between time and space.

Here we will concentrate on off-chip memory readout schemes and propose several different readout strategies in the next section. Finally, we will discuss how to optimize the readout scheme.

### 5.2 Off-Chip Memory Readout Schemes

It is noticed that "fired" and "unfired" are the only two states of the pixels at each sampling time instance. Thus, logic "1" and "0" can be used to describe these states. Therefore, the frame at one sampling time instance is a binary image that is suitable for some types of compression. The multiple sampling actually decomposes the multilevel image into a series of binary images. Each binary image can be compressed via one of several well-known binary compression methods, such as constant area coding, 1-D run-length coding (RLC), and predictive differential quantization [58]. Since raster scanning is adopted in our time-based CMOS imager design, a one-dimension compression method is applicable. One of the widely used compression methods is the 1-D RLC technique, which represents each row of a frame or bit plane by a sequence of lengths that describe successive runs of "fired" and "unfired" pixels and thus, it is not suitable for the time-based CMOS imagers because a counter is required to count the "runs" and all the pixels in a row will be scanned as that shown in literature [20], which seriously limits the readout speed. Here, we will first propose two new readout methods that only read out the necessary information of the "fired" pixels so that the readout speed can be dramatically increased compared with the conventional approach. These two methods are the absolute address coding (AAC) strategy and the relative address coding (RAC) strategy. The basic idea of each is to read out the addresses of the "fired" pixels at each

sampling time instance and the address information localizes the position of the corresponding pixel during the image reconstruction.

In the next sections, we will discussed the advantages and disadvantages of the two methods and then propose R&AAC readout scheme based on them. All the three readout schemes are serial processing schemes. Lastly, we will also propose a parallel readout scheme and compared it with the serial readout schemes.

**5.2.1 AAC Readout Scheme**

Absolute address coding scheme will read out exactly the addresses of those pixels that "fired". Figure 5.3 depicts the simplified diagram of the AAC readout scheme. The column addresses are wired in the D flip-flops (DFFs). When the "row select" signal is enabled, the digital outputs of the pixels are used as control signals to the DFFs. A logic "1" output will connect a DFF to the shift registers, while a logic "0" output will disconnect a DFF from the shift registers. When the shift registers are built up, the addresses will be shifted out of chip from the right to the left (i.e.: send out small addresses first). So, this is a parallel-in-serial-out (PISO) shift registers scheme. Let us use the row pixel vector of {0, <u>1</u>, <u>1</u>, 0, 0, <u>1</u>, 0, 0} as an example. The AAC scheme will read out the addresses of the underlined pixels, which are in the positions of 2, 3 and 6. In Figure 5.3 we can see that only those DFFs that store the addresses of the "fired" pixels are connected together to form the shift registers and these addresses will be read out, while the others are skipped.

Finally, when the end-mark is shifted out, the readout processing of the row completes. This end-mark helps distinguish row addresses from column addresses in the

received data (see Table 5.1) and indicates the end of a row data. Table 5.1 shows the structure of the readout data for a 32×32 pixel array with a total number of samples of $S$.

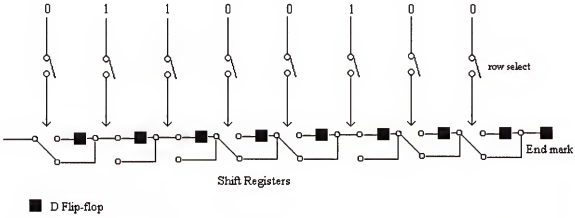



Figure 5.3 The simplified AAC diagram for the example of row pixel vector {0,1,1,0,0,1,0,0}.

Since the outputs of the AAC scheme are only the addresses of the "fired" pixels, the total number of readout pixels at one sample time instance is one bin of the sampled image histogram. Thus, the histogram of the sampled image can be achieved easily. Here we define the readout compression ratio (RCR) as

$$RCR = \frac{total \quad number \quad of \quad pixels \quad of \quad the \quad frames}{number \quad of \quad pixels \quad being \quad read \quad off \quad chip}$$

If there are 256 samplings, then the average RCR of one bit frame is 256. But the AAC readout scheme suffers from the problem of having to read out a big part of the whole bit frame at uniform luminance cases, which will dramatically increase the readout process load and reduce the readout speed. Actually, for the long successive "1" cases, we do not have to read out the pixels one by one, but we can implement the idea of the 1-D RLC to just read out the boundaries of the long sequence of "1". In the next section, we will discuss this idea in detail. The readout scheme based on this idea is named as relative address coding (RAC) scheme.

Table 5.1 The structure of the readout data for a 32×32 pixels array with a total number of samples of *S*.

| **Time index 0** |
|---|
| **Row1** |
| Column address 1 |
| Column address 2 |
| ... |
| Column address n1 |
| End mark |
| **Row 2** |
| Column address 1 |
| Column address 2 |
| ... |
| Column address n2 |
| End mark |
| **Row 3** |
| ... |
| End mark |
| **Row 32** |
| ... |
| End mark |
| **Time index 1** |
| **Row 1** |
| Column address 1 |
| ... |
| **Time index S−1** |
| ... |

### 5.2.2 RAC Readout Scheme

The idea of the RAC scheme is to read out the addresses of those pixels in each row of a bit frame where logic state change occurs. For example, for a pixel sequence of {0,1,1,0,0,1,0,0}, the logic states of the pixels at position of $1^{st}$, $2^{nd}$, $4^{th}$, $6^{th}$, and $7^{th}$ change from "1" to "0" or from "0" to "1", so the addresses of the $1^{st}$, $2^{nd}$, $4^{th}$, $6^{th}$, and $7^{th}$ pixel will be read out. A block diagram of a CMOS imager with the RAC readout scheme is similar to that in Figure 5.1 except that we should replace the AAC readout scheme with the RAC readout scheme. In practice, an extra logic value of "0" is placed at the head of the row vector to help determine the "1" and "0" order. Thus, the first address sent out is the position where the first "1" is located. The second address will be the

position where the first "0" located after the "1" sequence. The third address is the address of the first "1" located after the "0" sequence, and so on. For instance, for the row pixel vector of {0, 1, 1, 0, 0, 1, 0, 0}, the RAC scheme will read out the addresses of the underlined pixels, which are in the positions of 2, 4, 6 and 7.

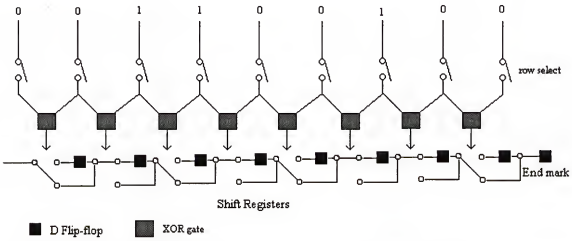



Figure 5.4 The simplified RAC diagram for the example of row pixel vector {0,1,1,0,0,1,0,0}.

Figure 5.4 depicts the simplified diagram of the RAC strategy. The adjacent digital outputs of the pixels in one row will perform the XOR logic. The output of the XOR gates will select the DFFs to form shift registers. A logic "1" output will connect a DFF to the shift registers, while a logic "0" output will disconnect a DFF from the shift registers. The built-up shift registers are also PISO shift registers that have the same working mechanism as those in the AAC readout scheme. An example of the connection in 1-bit shift registers is also depicted in Figure 5.5. The digital output of the row is the same one used in the section 5.2.1. When the "row select" signal is enabled, the XOR logic outputs will connect the $2^{nd}$, $4^{th}$, $6^{th}$, and $7^{th}$ DFFs to form shift registers as shown in Figure 5.5. We can see that only those DFFs that store the addresses of the pixels where

logic states change occurs are connected to form the shift registers and these addresses will be read out and the others are skipped. The difference between two adjacent readout addresses is the successive "run" length of "1" or "0".

It can be seen that the RAC scheme is to detect the boundaries of successive "1"s. The saved numbers of readout operations is different for different cases due to different "1" sequences. This is shown in table 5.2. It is clear that if "1"s are scattered among "0"s, more pixels will be picked than what is necessary. Since two pixels need to be picked to identify the boundary, for one "1" surrounded by "0"s case, the saving is -1, for two successive "1" case, there is not loss or saving, and for N successive "1" case where N is greater than 2, the earning is (N-2) and the larger the N is, the more the saving is. Therefore, the RAC scheme is efficient when a long successive run of "1"s exists in the bit frame. The worst case of RAC scheme is to read out the whole frame.

Table 5.2 The saved numbers of readout operations for the RAC readout scheme

| Number of successive "1" | Number of picked pixels | Number of saving |
|---|---|---|
| 1 | 2 | -1 |
| 2 | 2 | 0 |
| 3 | 2 | 1 |
| 4 | 2 | 2 |
| ... | ... | ... |
| N | 2 | N-2 |

While for high-resolution images (e.g. >=256 gray scales), the single "1" case occurs much more ofen than other cases in a bit frame. This results in that the RAC readout scheme will readout more pixels than necessary. Some experimental results are listed in Table 5.3. All the images in this table are represented with 8-bits representation. These experimental results indicate that the single "1" case occurs more often than other cases.

Table 5.3 The simulation results for normalized readout pixels for the RAC readout scheme

| Image | Image size | Dynamic range (dB) | $\frac{readout\ \ pixel\ \ number}{image\ \ size}$ |
|---|---|---|---|
| lena | 256×256 | 48 | 1.68 |
| cameraman | 256×256 | 48 | 1.69 |
| F16 plane | 256×256 | 48 | 1.78 |
| trevor | 256×256 | 48 | 1.80 |
| sunset | 480×720 | 97 | 0.63 |
| suntree | 480×720 | 103 | 1.74 |
| church | 768×512 | 113 | 1.49 |
| suntree2 | 480×720 | 124 | 1.70 |
| window | 480×720 | 130 | 1.60 |
| church2 | 480×720 | 168 | 1.49 |

In order to remove the redundancy further, the RAC readout scheme and the AAC readout scheme can be combined together. That is described in the next section.

### 5.2.3 R&AAC Readout Scheme

The RAC readout scheme has the advantage of removing redundancy when successive runs of "1"s are not less than 2 (see Table 5.2). While, the AAC readout scheme is very suitable when a single "1" is embedded in "0"s. The combination of these two methods, the relative and absolute address coding (R&AAC) readout scheme, will combine the advantages of each method and compensate their disadvantages. The "010" pattern, which denotes the single "1" case, is the one where the AAC readout scheme should be implemented. For other patterns, the RAC readout scheme should be implemented. The block diagram of a CMOS imager with the R&AAC readout scheme is similar to that in Figure 5.1 except that the AAC readout scheme is replaced by the R&AAC readout scheme. A simplified diagram of the R&AAC scheme is shown in Figure 5.5.

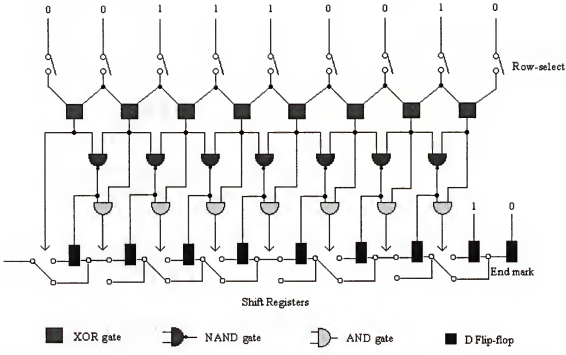



Figure 5.5 The simplified diagram of the R&AAC readout scheme. The first bit in the example is the extra logic value "0" as that used in the RAC readout scheme. The first row DFFs in the Figure is used to store the distinction bits. The distinction bits for RAC and AAC are 1 and 0 respectively.

The example shown in this Figure is a row pixel vector of {0, $\underline{1}$, 1, 1, $\underline{0}$, 0, $\underline{1}$, 0}. We also need to add an extra logic bit "0" in front of the row vector as that used in the RAC readout scheme. The R&AAC readout scheme will read out the addresses of the underlined pixels, which are in the positions of 2, 5 and 7. The XOR gates are used to detect the logic states changes of the row pixels. To combine with the AAC readout scheme, the "010" pattern should be separated from other patterns. Notice that only the pattern of "010" has successive "1" XOR output. For example, the XOR gates outputs of the row pixel vector {0, 1, 1, 1, 0, 0, 1, 0} are {0, 1, 0, 0, 1, 0, 1, 1}. Therefore, a two-input NAND gate can be used to represent this information. Finally, the XOR gates outputs and the NAND gates outputs pass through the AND gates to work as the control signals of the DFFs. A logic "1" output of the AND gates will connect the corresponding

DFF to the shift registers and a logic "0" output of the AND gates will disconnect the corresponding DFF from the shift registers. To distinguish the RAC readout scheme output from the AAC readout scheme output, we need to add one more bit named as the distinction bit (see the first row of DFFs in the shift registers shown in Figure 5.5). Here "1" is used to indicate that the address is a RAC readout scheme output, "0" is used to indicate that the address is an AAC readout scheme output. The distinction bit of the last row pixel is set to 1, and the distinction bit of the end mark is set to 0 so that there do not exist two addresses that have same codes. In summary, the XOR gates fulfill the RAC scheme, the NAND gates fulfill the AAC scheme, and lastly the AND gates combine these two schemes together.

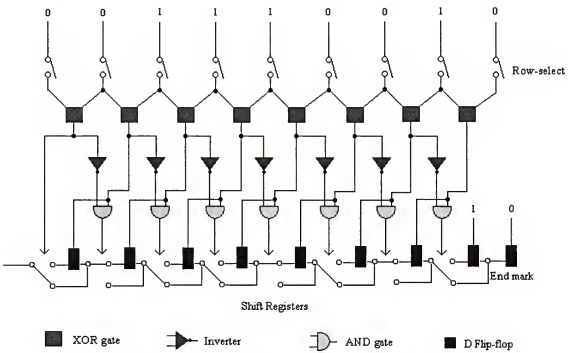



Figure 5.6 A simplified version of Figure 5.5.

Observing Figure 5.5, we see that each R&AAC readout scheme output (the output at the AND gates) is the logic output of three successive inputs. Only the patterns "001" and "110" have logic "1" outputs. Thus, the R&AAC logic can be described by

$$\overline{(a \otimes b)(b \otimes c)}(b \otimes c) \tag{5.1}$$

$$= \overline{(a \otimes b)}(b \otimes c) \tag{5.2}$$

The formula (5.1) is the logic expression of Figure 5.5, and it can be simplified to formula (5.2). Therefore, inverters can be used to replace the NAND gates in Figure 5.5 and the circuit simplifies as that illustrated in Figure 5.6.

Table 5.4 The simulation results for normalized readout pixels for the R&AAC readout scheme

| Image | Image size | Dynamic range (dB) | $\frac{readout\ \ pixel\ \ number}{image\ \ size}$ |
|---|---|---|---|
| lena | 256×256 | 48 | 0.93 |
| cameraman | 256×256 | 48 | 0.87 |
| F16 plane | 256×256 | 48 | 0.91 |
| trevor | 256×256 | 48 | 0.92 |
| sunset | 480×720 | 97 | 0.39 |
| suntree | 480×720 | 103 | 0.92 |
| church | 768×512 | 113 | 0.83 |
| suntree2 | 480×720 | 124 | 0.91 |
| window | 480×720 | 130 | 0.87 |
| church2 | 480×720 | 168 | 0.84 |

Table 5.4 shows the experiment results of implementing the R&AAC readout scheme on the images in table 5.3. It is clear that the normalized readout pixel number is less than 1, which is consistent with what we have discussed. The R&AAC readout scheme has two more gate delays than the RAC readout scheme and three more gate delays than the AAC readout scheme. However, the R&AAC readout scheme provides more compression in the readout than either the RAC or the AAC readout scheme. On the other hand, the encoding and decoding for the R&AAC readout scheme are the most complex ones as the methods with the AAC readout scheme being the simplest. All three readout schemes use PISO shift registers as the output engine, leading to bit serial outputs. Another way to quickly readout data is to use as many parallel outputs as

possible. We will propose a parallel readout scheme in the next section and discuss which scheme provides the fastest readout under same clock frequency.

### 5.2.4 Parallel Readout Scheme

A parallel readout scheme is proposed here as shown in Figure 5.7. This scheme adopts the idea of on-chip memory that uses the output of each pixel as the write-enable signal to enable the "write" processing of the time index to the memory (see section 5.1), but here the memory is an off-chip memory. Multiplexers are used since we may not be able to send out the whole row simultaneously due to the finite pin number of a package. It is obvious that the more pins we can use, the more bits we can read out at one time and the faster we can read data out. Multiplexing method was also implemented in Kishore [20], but here, the off-chip memory is used to store the sampling time index instead of the bit plane. This makes the off-chip memory much smaller.

To compare the advantages and disadvantages of the serial readout schemes and the parallel readout scheme, we will use the following simple model. Assume we have an $N \times N$ pixel array and the average percentage of readout pixels in each row is $y\%$ for a serial readout scheme. The maximal available parallel data output pins for the parallel readout scheme is $M$ ($M<=N$). Then the total numbers of readout operations in one bit frame for a serial readout scheme is

$$N_{serial} = N^2 y\% \tag{5.3}$$

Then the total numbers of readout operations in one bit frame for the parallel readout scheme is

$$N_{parallel} = \frac{N^2}{M} \tag{5.4}$$

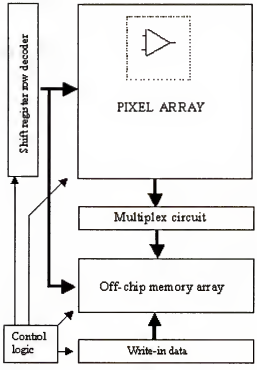



Figure 5.7 Block diagram of the parallel readout scheme.

If we ignore the difference of the switching times of the two types of readout schemes, the relation of $y\%$ and $\frac{1}{M}$ will determine which scheme has a shorter readout time in the reading of one bit frame. In general, $y\%$ will be around 2% as shown in VanRullen et. al. [55], which corresponds to an $M$ value of about 50. Thus, if the maximal data output pins are less than 50, the serial readout schemes will have a shorter readout time. In addition, only $\log_2 N + 1$ data output pins are needed in a serial readout scheme, which is smaller than 50 generally. These two advantages make the serial readout schemes more competitive. Also notice that the data output pin number for a serial readout scheme just increases 1 for every doubling of the pixel array size.

### 5.3 Conclusions And Discussions

We have proposed several different readout schemes, discussed their applications and compared their advantages and disadvantages. The serial readout schemes have more compression in readout than the multiplexed readout scheme generally. In the serial readout schemes, the R&AAC readout scheme has the best readout compression ratio but the most complex circuit realization and image reconstruction process and more latency. The AAC readout scheme has a moderate readout compression ratio and the simplest circuit realization and image reconstruction process, but suffers from the worst case of uniform luminance that requires the whole frame to be read out within a short interval. This worst case is similar as that described in literature [19]. The RAC readout scheme can successfully solve this problem of the worst case and therefore, the R&AAC readout scheme that combines the RAC and AAC readout schemes together does not have this problem and has the advantages of both the AAC and RAC readout schemes. If the complexity and small latency can be tolerated, the R&AAC readout scheme is the outstanding one.

Row parallel processing can be combined with the serial readout schemes. $M$ rows in parallel processing will increase the throughout speed by $M$ times. The drawback is the output data pins will also increase $M$ times.

As we know, for those rows that have many pixel addresses to read out, it will need longer time, but for those rows that have few or even no pixel addresses to read out, it will need less time. If the same readout time is assigned for each row, the disadvantage is obvious that the time may be not enough for some rows and too much for other rows.

To solve this problem, an adaptive scheme can be used to provide different time for different rows based on the load of the row. Thus, we will not assign the same readout time to the rows as what the conventional raster scan does. Instead, in the sampling of one bit frame, as long as one row is read out, the end mark of this row will start the sampling of the next row. The sampling time for one particular row is also different for different images. Thus, it is actually signal dependent. So the combination of the adaptive scheme with the serial readout schemes can provide more flexibility and is the optimal readout scheme. A summary of the readout schemes for digital imagers is illustrated in Figure 5.8.

In the next chapter, we will present the circuit design for the time-based CMOS imager system.

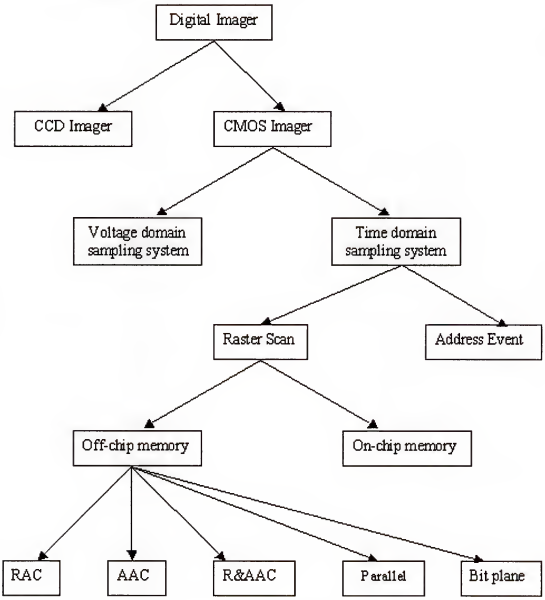



Figure 5.8 Readout schemes for digital imagers

# CHAPTER 6
# CIRCUIT DESIGN

We have solved the problem of how to achieve ultra-wide dynamic range imaging for a CMOS imager and proposed the optimal readout scheme for off-chip memory. In this chapter, we will discuss the circuit realization of an ultra-wide dynamic range CMOS imager. A novel DPS pixel design is proposed. The comparator design and the disable circuit design in the pixel are discussed in detail. Then we present the readout system design. Finally, the generation of timing table and reference voltage table are discussed and the whole system level architecture is presented.

## 6.1 Pixel Design

As we have discussed in Chapter 3, digital pixel sensors are required for time-based CMOS imagers. The incident light discharges the reset photodiode and causes the photodiode voltage $V_{ph}$ to decrease. When the $V_{ph}$ drops to the value of reference voltage, the comparator will toggle. The output of a comparator is directly a 1-bit digital output. Since time-based CMOS imagers detect only those pixels that "fire" at the current sampling time instance but not those that have "fired" at the previous sampling time instance, the "fired" pixels should be reset to "unfired" status and kept in this state until the whole sampling processing is finished. Here we call this reset function as the disable function. Thus, it is clear that there are three important parts in one pixel: photodiode, comparator and disable circuit. The block diagram of the DPS pixel is shown in Figure 6.1.

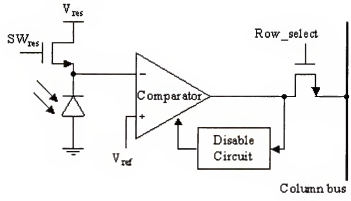



Figure 6.1 Block diagram of the DPS pixel.

In the Figure, $V_{res}$ is the reset voltage. To avoid the effect of transistor threshold voltage variation, $V_{res}$ is selected carefully so that $V_{dd} - V_{res}$ is greater than the average reported threshold voltage and $V_{res}$ is large enough to keep the dynamic range of $V_{ref}$ high. Here $V_{dd}$ is 5V, and $V_{res}$ is selected to be 3V. Another way is to use a PMOS as the reset transistor. In this case, the reset voltage can be chosen to be $V_{dd}$ and threshold voltage variation can be avoided. When $SW_{res}$ is equal to $V_{dd}$, the photodiode is reset to $V_{res}$. When it becomes low (ground), the photodiode voltage $V_{ph}$ decreases. When $V_{ph}$ drops to below $V_{ref}$, the comparator toggles. The row_select signal will connect the comparator output to the column bus when it is high, or disconnect the comparator output from the column bus when it is low. Since time-based CMOS imagers detect the pixels that are "fired" at the current sampling instance, but are "unfired" at the previous sampling instance, the "fired" pixels should be reset to the "unfired' state and kept in that state until the end of the image sampling. This is fulfilled by the disabling circuit shown in Figure 6.1. The designs of the two major components: comparator and disabling circuits are discussed in detail in the next sections.

## 6.2 Comparator Design

To optimize the performance of a comparator, one has to consider the resolution, speed and power dissipation trade-off (in a situation of reasonable devices size, i.e. offset considerations). To achieve high resolution, large gain structures are necessary. To achieve high speed processing, wide unit frequency response is required and this indicates a relatively small gain. Possibly the most power-efficient high-speed design method involves combining a low-gain (high bandwidth) stage with a positive feedback track-and-latch circuit.

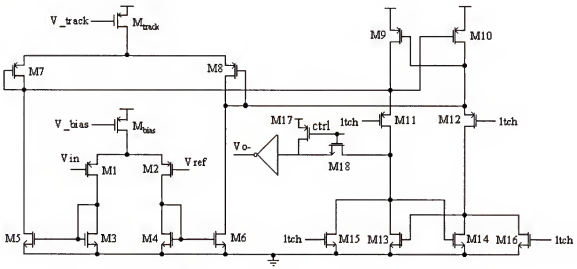



Figure 6.2 Circuit diagram of the comparator. Power supply is 5V.

In this section we will present a new approach to the design of a CMOS differential latched comparator. The core of the comparator is an amplifier/latch combination as shown in Figure 6.2. When the "V_track" $\left(V_track = \overline{ltch}\right)$ pulse is low, that is the comparator works in track mode, the circuit works as a pre-amplifier with a gain of about 11.2dB because the diode-connected transistors $M_7$ and $M_8$ are fully turned on. The combination of the diode-connected transistors of the gain stage and the transistors of the positive-feedback loop acts as moderately large impedance and gives

gain from the preamplifier stage to the track-and-latch stage. The diode-connected loads of the preamplifier stage give a relatively small gain in order to maximize speed, while still buffering the kickback from the input circuitry.

### 6.2.1 Gain Of The Preamplifier

The gain $A_v$ of the preamplifier can be calculated as following.

$$A_v = A_{v1} \cdot A_{v2} \tag{6.1}$$

where

$$A_{v1} = g_{m1} \cdot R_{o1} = g_{m1} \cdot \left( \frac{1}{g_{m3}} // r_{o1} \right) \approx \frac{g_{m1}}{g_{m3}} \tag{6.2}$$

$$A_{v2} = g_{m5} \cdot R_{o5} = g_{m5} \cdot \left( r_{05} // R_{o7} \right) = g_{m5} \cdot \left( r_{o5} // \frac{1}{g_{m7}} // R_L \right) \tag{6.3}$$

Here, $R_L$ is the resistance looking into the positive feedback loop. The open loop gain of the feedback is

$$A_{oL} = g_{m10} \cdot \left( \frac{1}{g_{m8}} // r_{o6} \right) \cdot g_{m9} \cdot r_{o9} \approx g_{m10} \cdot \frac{1}{g_{m8}} \cdot g_{m9} \cdot r_{o9} \tag{6.4}$$

The positive feedback at the output port of the loop is the shunt feedback, so the equivalent output resistance is

$$R_L = -\frac{r_{o9}}{A_{oL}} = -\frac{g_{m8}}{g_{m9} \cdot g_{m10}} \tag{6.5}$$

Substituting equations (6.2) ~ (6.5) to equation (6.1), we can obtain the gain of the preamplifier as

$$A_v \approx \frac{g_{m1} g_{m5} g_{m8}}{g_{m3} \left( g_{m7} g_{m8} - g_{m9} g_{m10} \right)} \tag{6.6}$$

If the denominator approaches 0, a very high gain can be achieved. To optimize the speed, we only need a relatively small gain. Here the bias voltage is set to be 4.25V, the preamplifier provides a DC gain of 11.2dB and cutoff frequency of 53.2MHz.

**6.2.2. Regenerative Latch**

In the latch phase, the "ltch" pulse goes low and "V_track" pulse goes high, the diode-connected transistors $M_7$ and $M_8$ are disconnected from the output nodes, and the amplifier becomes a regenerative latch.

The regenerative output latch is based on an approach presented in literature [59]. The latch consists of discharge transistors M15 and M16, an n-channel flip-flop ($M_{13}$ and $M_{14}$), a p-channel flip-flop ($M_9$ and $M_{10}$), and a pair of p-channel transmission gates for strobing ($M_{11}$ and $M_{12}$). The output is taken at the drain terminals of $M_{11}$ and $M_{12}$ instead of the corresponding source nodes in order to increase the regeneration speed and reduce the offset as illustrated in Deen et al. [60]. In addition, the strobing transistors isolate the n-channel flip-flop transistors from the preamplifier output nodes. The transistors in the latch are minimum size transistors to increase the latch speed. Since the comparator output is isolated from the input by the current mirrors, there is no kickback effect, and a full logic swing is obtained from the comparator output. In addition, the "ltch" pulse will reset the comparator output after each decision cycle and therefore, there is no hysteresis in the design.

Because the transistors $M_7$ and $M_8$ are diode-connected, the output voltage of the preamplifier is limited between $(V_{dd} - V_{th})$ to $V_{dd}$ where $V_{th}$ is the threshold voltage of the transistor. This voltage will not cause the transistors of the p-channel flip-flop to open simultaneously in the track mode and therefore, guarantees the correct operation. The

positive feedback regenerative latch will amplify the preamplifier output further with very high gain as pointed out by Razavi et. al. [61] and give the digital output directly.

The time constant of the latch, when it is in latch phase, is estimated roughly as

$$\tau_{ltch} = K\frac{L^2}{\mu_n V_{eff}} \tag{6.7}$$

where $K$ is a constant between 2 an 4 as shown in Johns et al. [62]. $L$ is the channel length of the transistors, $\mu_n$ is the mobility of electron and $V_{eff}$ is the effective gate-source voltage. This formula implies that $\tau_{ltch}$ depends primarily on the technology and not on the design (assuming a reasonable design is used that maximizes $V_{eff}$ and minimizes load capacitance). If the voltage difference of $\Delta V_{digit}$ is necessary to be obtained for succeeding logic circuitry to safely recognize the correct output value, then the time necessary for this to happen is given by

$$\tau_{ltch} = K\frac{L^2}{\mu_n V_{eff}}\ln\left(\frac{\Delta V_{digit}}{\Delta V_0}\right) \tag{6.8}$$

where $\Delta V_0$ is the initial difference between the two input voltages to the latch. Table 6.1 shows the simulated latch time for different initial difference between two input voltages. Here, $\Delta V_{digit}$ is 5V.

Table 6.1 The simulated latch time for different initial difference between input voltages

| Initial difference between input voltages | Simulated latch time (ns) |
|---|---|
| 10mV | 1.5 |
| 1mV | 2.1 |
| 0.1mV | 2.8 |
| 10μV | 3.2 |

At a sampling time instance, if $\Delta V_0$ is very small, the latch time can be large (this is called metastability). The output voltage of the latch may not increase enough to be recognized as the correct logic value by succeeding circuitry. So a sufficient time needs

to be allocated to allow the latch to reach its steady state. In our design 3ns is adopted as the time for the latch to reach steady state. Within this time, the gain of the latch is able to detect 0.1mV difference of the inputs.

### 6.2.3. Common-Mode Input Voltage Of The Comparator

Simulated results show that the common-mode input voltage can be as low as 0V (differential PMOS input pair) and as high as 4.1V. The gain of the preamplifier in this range varies from 9.9dB to 11.6dB, a relatively stable gain.

### 6.2.4. Offset Of The Comparator

The offset of a comparator includes systematic offset and offset caused by mismatch during manufacturing. The offset voltage in the latch circuit is systematically minimized by the latch topology [63]. The main source of offset is in the differential pair. There are two kinds of offset sources in a MOS differential pair. They are charge density fluctuation, such as surface states and impurity doping concentration, and dimension fluctuation.

The offset associated with the differential input stage (PMOS input transistors and NMOS load transistors) is given by

$$V_{os} = D_p \cdot \left( \frac{1}{L_p} + \frac{1}{L_n} \right) \cdot \sqrt{\frac{I_{bp}}{2k_p}} + \Delta V_{tp} + \sqrt{\frac{k_n}{k_p}} \cdot \Delta V_{tn} \tag{6.9}$$

where $D_p$ denotes the channel length mismatch between the input transistors pairs, $L_p$ and $L_n$ are the PMOS an NMOS channel length respectively, $I_{bp}$ is the biasing current of the input pair, $\Delta V_{tp}$ and $\Delta V_{tn}$ are the threshold voltage variations for p-type and n-type MOS transistors respectively [64]. The first term comes from the dimension fluctuation in the differential input pair and load transistors. It tells us that longer channel input transistors

are helpful in reducing offset. However, we need to remember that a too long channel length will cause transistor size too big and large parasitic capacitance that reduces the circuit speed. A small bias current contributes to the reduction of offset too. In this design, the bias current is small ($I_{bp}$=0.35uA for each branch). Also using layout technologies, such as common centroid configurations, the transistor mismatch caused by gradients can be greatly reduced. The second and third terms represent the contribution from the charge fluctuation. Because the surface state density and impurity concentration can be well controlled, these terms are insignificant compared to the first. Simulation results indicate that the offset is 2.8mV for every 1% mismatch between the channel lengths and 1.0mV for every 1% mismatch between the transistor widths.

### 6.2.5. Settling Time Of the Comparator

The settling time is defined to be the time it takes for an amplifier to reach a specified percentage of its final value when a step input is applied. It consists of two distinct segments: linear and nonlinear settling time segments. The linear settling time segment is due to the finite unity-gain frequency of the amplifier, and thus it sets a minimal value for the overall settling time independent of the step size of the amplifier's output. The nonlinear settling time segment is due to the slew-rate limiting, and thus it is strongly dependent on the output step size. In the CMOS imager case, the input photodiode voltage is linearly decreasing with a DC sweep. When the slope of the input signal is very steep, the amplifier may reach the slew-rate limitation. When the slope of the input signal is small, the amplifier will not reach the slew-rate limitation and the settling time is decided by the unity-gain frequency. The unity-gain frequency for the amplifier to settle is estimated within the linear settling–time segment. Since the positive

feedback in the latch mode provides very high gain and wide bandwidth, the limitation of bandwidth is in the preamplifier. So a sufficient large cutoff frequency of the preamplifier is required.

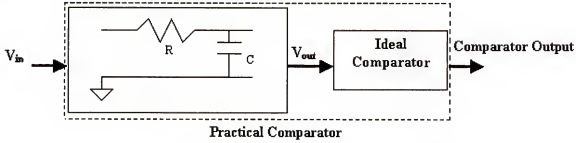



Figure 6.3 A simplified comparator model. It consists of a first-order RC circuit part and an ideal comparator part. The RC circuit provides the dominant pole and the ideal comparator has infinite bandwidth.

Table 6.2 The ac analysis results of the preamplifier.

| Vbias (V) | Ibias (uA) | DC gain (dB) | $f_c$ (MHz) | 1st pole phase margin |
|---|---|---|---|---|
| 4.30 | 0.13 | 10.3 | 26.1 | 80.0 |
| 4.25 | 0.70 | 11.0 | 53.2 | 67.5 |
| 4.20 | 0.88 | 11.6 | 91.5 | 62.7 |
| 4.15 | 2.04 | 12.0 | 130.5 | 61.9 |
| 4.10 | 4.22 | 12.2 | 169.2 | 63.1 |

To estimate the transient response of a comparator, we use a simplified first-order comparator model as shown in Figure 6.3, which has a 90-degrees phase margin. In table 6.2 we list the AC parameters of the comparator preamplifier working in different bias current. The phase margin of the preamplifier is around 62-degrees most of the time, which indicates that the second pole is not very far away from the dominant pole. In the first order approach, the simplified model shown in Figure 6.3 is still reasonable.

According to the Kirchhoff's current law, the differential equation for the above circuit is described by

$$\frac{dV_{out}(t)}{dt} + \frac{V_{out}(t)}{\tau} = \frac{V_{in}(t)}{\tau} \tag{6.10}$$

Where $\tau = RC$. The solution to the first-order differential equation is

$$V_{out}(t) = e^{-\frac{t}{\tau}}\left[C_1 + \int \frac{V_{in}(t)}{\tau} e^{\frac{t}{\tau}} dt\right] \tag{6.11}$$

where $C_1$ is a constant that is decided by the initial conditions. In the photodiode sensor application, the input voltage to the comparator is the photodiode voltage, which is

$$V_{in}(t) = V_{res} - \frac{I_{ph} t}{C_{ph}} \tag{6.12}$$

where $C_{ph}$ is the capacitance at the photodiode cathode, $V_{res}$ is the reset voltage and $I_{ph}$ is the photocurrent. Let's use

$$k = \frac{I_{ph}}{C_{ph}} \tag{6.13}$$

to represent the slope of the DC sweep of the input voltage. Assume $C_{ph}$=5fF, and $I_{ph}$ has the range of 0.5fA to 0.5nA, the $k$ value is from 0.1V/us to 0.1uV/us. Substituting equations (6.12) and (6.13) into the equation (6.11), the following result is obtained.

$$\begin{aligned} V_{out}(t) &= e^{-\frac{t}{\tau}}\left[C_1 + \int \frac{V_{res} - kt}{\tau} e^{\frac{t}{\tau}} dt\right] \\ &= e^{-\frac{t}{\tau}}\left[C_1 + \int \frac{V_{res}}{\tau} e^{\frac{t}{\tau}} dt - k\int \frac{t}{\tau} e^{\frac{t}{\tau}} dt\right] \end{aligned} \tag{6.14}$$

Let $y = \frac{t}{\tau}$. Since

$$\begin{aligned} \int y e^y dy &= \int y de^y \\ &= ye^y - \int e^y dy \\ &= ye^y - e^y \end{aligned}$$

we can obtain that

$$\int \frac{t}{\tau} e^{\frac{t}{\tau}} dt = \tau \left( \frac{t}{\tau} e^{\frac{t}{\tau}} - e^{\frac{t}{\tau}} \right) \tag{6.15}$$

Substituting equation (6.15) into equation (6.14), we obtain that

$$V_{out}(t) = e^{-\frac{t}{\tau}} \left[ C_1 + V_{res} e^{\frac{t}{\tau}} - k e^{\frac{t}{\tau}} (t - \tau) \right]$$

$$= C_1 e^{-\frac{t}{\tau}} + V_{res} - k(t - \tau) \tag{6.16}$$

The $V_{out}$ consists of two parts. The first part is the first term in equation (6.16), which is an exponentially decreasing term. The second part is the second and third terms in equation (6.16), which is $\tau$ delay of the input $V_{in}$. This delay is exactly the time constant of the comparator and can be described as

$$\tau = \frac{1}{\varpi_{-3dB}} = \frac{1}{2\pi f_{-3dB}} \tag{6.17}$$

In the CMOS imager application,

$$V_{out}(0) = V_{res} \tag{6.18}$$

This initial condition gives that

$$C_1 = -k\tau \tag{6.19}$$

Thus, equation (6.16) becomes

$$V_{out}(t) = -k\tau e^{-\frac{t}{\tau}} + V_{res} - k(t - \tau)$$

$$= V_{in}(t) + k\tau \left( 1 - e^{-\frac{t}{\tau}} \right) \tag{6.20}$$

For different incident light intensities, $k$ values (DC sweep slopes) are different. Thus, the difference between $V_{out}$ and $V_{in}$ are different for different pixels with different incident light intensities at same sampling time. Bright pixels have larger error than dark

pixels. To achieve an accurate comparison result, we need to let $V_{out}$ be as close as possible to $V_{in}$. There are two ways that can help realize this target. The first is to sample the pixels $\tau$ time later than the theoretical sample time. Another way is to make $\tau$ as small as possible. A small $\tau$ will cause the exponentially decreasing term to decrease to zero very fast (e.g. when $t > 5 \cdot \tau$, $e^{-\frac{t}{\tau}} < 0.007$). The more important is that the second term in equation (6.20) will become much smaller than the first term quickly. Since the model is a simplified one and a precise $\tau$ is hard to measure for an imager pixel array, the method of delaying sampling time by $\tau$ is infeasible. But $\tau$ can be design to be a very small value to improve the comparison accuracy as much as possible. The worst case exists in the high light intensity range, where $k$ value is big (e.g 0.1V/us). For example, the time constants corresponding to –3dB frequencies of 200Hz, 2kHz, 20kHz, 200kHz, 2MHz, 20MHz are 0.8ms, 80us, 8us, 0.8us, 80ns, 8ns respectively. If $k$=0.1V/us, $V_{res}$=2.5V, the relation of $V_{out}$ versus $V_{in}$ is shown in Figure 6.4. It is clear that $V_{out}$ follows $V_{in}$ well for small $\tau$.

In our design, we decide that the delay time to be no greater than 3ns or cutoff frequency of the preamplifier to be greater than 53MHz. Thus, the settling time is at most 13.8ns for 1% accuracy. Our simulation results (see table 6.2) show that this is not hard to achieve. Here, use the bias current of 0.7uA for the circuit.

**6.2.6 Discussions**

P-channel input transistors are used here to extend the common mode voltage and avoid charge trapping because p-channel transistors exhibit much less hystersis than n-channel transistors. Also the 1/f noise of p-channel transistors is less than the n-channel

transistors with the same size. The simulation results show that the input referred noise is about $92nV/\sqrt{Hz}$.

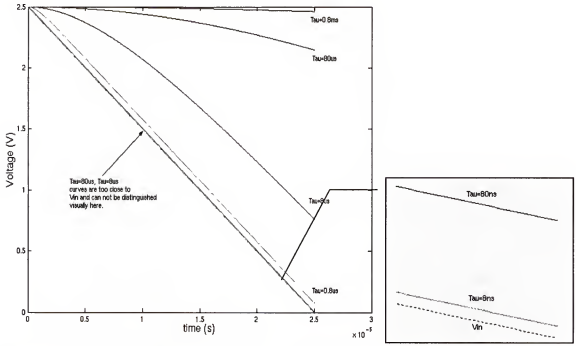



Figure 6.4 Relation of $V_{out}$ versus $V_{in}$ for the bright pixels ($k$=0.1V/us) in the simplified comparator model. Here $V_{res}$=2.5V, $\tau$ values are 0.8ms, 80us, 8us, 0.8us, 80ns, 8ns as shown in the Figure. It is clear that $V_{out}$ is very close to $V_{in}$ for small $\tau$.

### 6.3 Disable Circuit Design

In order not to lose any information, a fired pixel should be acknowledged before it is disabled. Thus, the disable processing should be assigned after the row has been selected and information of the "fired" pixels in this row has been read out. Also, in order to keep the "unfired" state until the end of the whole sampling, a latch should be used to store the disable state. The disable circuit in our design is shown in Figure 6.6. Figure 6.5(a) shows the block diagram of the disable processing. Two row shift registers (RSR) are implemented. The shift registers of the "row_select" form the row decoder and the shift registers of the "dis_e" form the row disable operation selector. The bit values of

the shift registers are wired in at the beginning of each sampling instance. These two set of shift registers are controlled by the same clock named as the RSR_clk.

Figure 6.5(b) shows the disable circuit inside a pixel. At the start of sampling, $SW_{res}$ is high and transistor $M_{dis}$ is closed. During the sampling, $SW_{res}$ is kept low. The "dis_e" signal in Figure 6.5(b) is one clock cycle delay to the "row_select" signal to guarantee the pixel is first detected before it is disconnected. The $SW_{res}$ is the reset control signal. At the beginning of a sampling, it is set to $V_{dd}$ to reset the photodiode and reset the ctrl signal at Figure 6.5(b) to be logical 1 such that the comparator is connected to the inverter at the output. When "row_select" is logical 1, "dis_e" is logic 0. At the next clock, "row_select" becomes logical 0 and "dis_e" becomes logical 1. If the comparator toggles, the output signal $V_{o-}$ becomes logical 1. Then, the disconnect circuit will force the "ctrl" signal to change to logical 0 and bypass the $V_{dd}$ to the inverter at the output such that the output of the pixel will always be logic 0. The "ctrl" signal can only be reset at the next sampling.

Note that both sets of shift registers have one more row than the pixel array. The "dis_e" shift registers are one row higher than the pixel array, and the "row_select" shift registers are one row lower than the pixel array. This arrangement guarantees that no rows of the pixel array will be missed from the disable processing. The last bit in the "row_select" shift registers is named as the end_frame bit. When it is "0", the pixel array is in latch mode and the raster scan is in process. When it becomes "1", the pixel array is in track mode and the raster scan is finished (the current frame has been sampled and the new sampling can start). Thus, it is clear that latch signal satisfies:

$$ltch = end_frame_bit$$

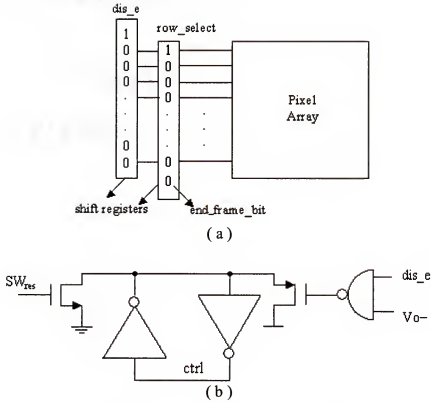



Figure 6.5 Disabling circuit diagram. (a) Block diagram of the disable processing. The "row_select" shift registers form the row decoder. The "dis_e" shift registers form the row disable operation selector. These two shift registers are controlled by same clock. (b) The schematic of disable circuit in one pixel. Vo- is the output of comparator.

### 6.4 Readout System Design

Figure 6.6 shows the block diagram of readout system with the R&AAC readout scheme. In our design, the whole pixel array resets and then enters the integration mode simultaneously. The frame is latched by the "sample" signal at the beginning of a sample. Then the raster scan method is used to scan the rows one by one and the column addresses of those "fired" pixels are read out. In order to detect the end of the readout of a row, an end mark is placed at the end of the column shift registers (CSR). A unique end mark should be selected to distinguish it from the true addresses. A logic circuit named "check logic" will detect the end mark. Only when the end mark is detected by the check logic, the logic "1" is output by the check logic circuit. Otherwise, the logic "0" is output.

Thus, the output of the check logic circuit can be used as the clock (RSR_clk) of the row shift registers. The clock of the CSR is named as the CSR_clk. When the pixel array is in the latch mode, the clock signal is passed and works as the CSR_clk. When the pixel array is in track mode, the clock signal is blocked and there is no clock signal for the CSR. The generation of the RSR_clk and the timing relations between the CSR_clk and the RSR_clk are shown in Figure 6.7.

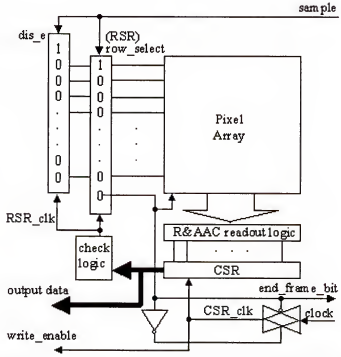



Figure 6.6 Block diagram of readout system with the R&AAC readout scheme.

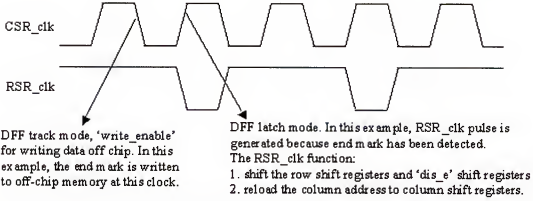



Figure 6.7 An example for the timing relations between the CSR_clk and the RSR_clk.

In summary, the functions of the end_frame bit signal are:

1) enable the last row to enter the disable processing
2) disconnect the CSR_clk from the clock signal when it is "1", otherwise, connect the CSR_clk to the clock signal
3) force the whole pixel array to enter track mode when it is "1", otherwise, keep the array in latch mode
4) enable the input of the next reference voltage

The functions of the "sample" signal are:

1) latch the bit frame
2) initialize the CSR and RSR
3) connect the CSR_clk to the clock signal

The functions of the "CSR_clk" signal are:

1) work as the clock of the CSR, up_down edge for track mode and down_up edge for latch mode
2) work as write enable signal of the off-chip memory

The functions of the "RSR_clk" signal are:

1) work as the clock of the RSR, up_down edge for track mode and down_up edge for latch mode
2) initialize the column addresses to the CSR

### 6.5 The Generation Of Referene Voltages And The "Sample" Signal

The diagram in Figure 6.8 illustrates the block diagram of the generation of the reference voltages and the "sample" signal. The $SW_{res}$ signal will reset the counter at the very beginning of the image capture.

A digital-to-analog conversion (DAC) is used to generate the reference voltages. The digital inputs of the DAC are stored in an off-chip memory ($V_{ref}$ table). The read processing of the $V_{ref}$ table is controlled by the end_frame bit signal.

A counter is used to decide the sampling time instance. The output of the counter is compared with the value in the timing table, which is also an off-chip memory. The read processing of timing table is also controlled by the end_frame bit signal.

### 6.6 The System Architecture Of The CMOS Imager

Based on the above discussions, the final system level architecture of the CMOS imager is shown in Figure 6.9. The information of the captured image is stored in the off-chip memory. Then it is read to a PC through the enhanced parallel port (EPP) interface. In the next chapter, we will analyze the noise in the system based on the circuit design presented in this chapter.

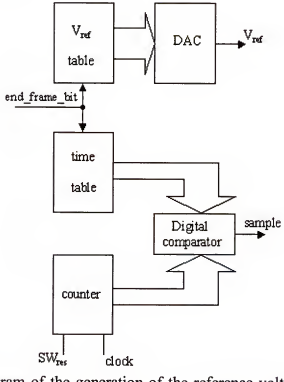



Figure 6.8 Block diagram of the generation of the reference voltages and the "sample" signal.

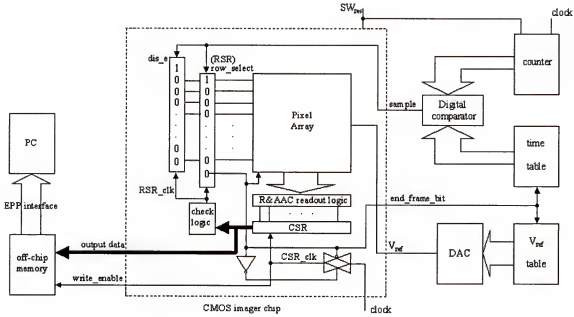



Figure 6.9 System level architecture of the time-based CMOS imager.

# CHAPTER 7
# NOISE ANALYSIS

In this chapter we will present the noise analysis for the proposed CMOS imager. This includes the temporal noise and the fixed pattern noise (FPN), which is an important type of non-uniformity in the pixel array. We also briefly discuss the role of image defects.

## 7.1 Temporal Noise

Temporal noise for a CMOS imager is the temporal variation in pixel output values under uniform illumination due to device noise, for example, thermal noise, shot noise, substrate noise, and supply voltage fluctuations. It increases with photocurrent (e.g. shot noise), but its effect is most pronounced at low signal levels (e.g. low luminance). We will see later that the signal to noise ratio (SNR) increases with the signal. Therefore, temporal noise under dark conditions ($I_{ph} = 0$) limits the minimum light intensity that can be measured. This will be discussed in detail in section 7.2. The digital pixel sensor (DPS) is used in our design. The comparator output works as the output of a one-bit ADC. Only the front end of the pixel that includes the reset transistor, photodiode, and preamplifier is implemented with analog circuits. Since the remaining parts of the pixel are digital circuits, the temporal noise is only an issue for the front-end analog components. Figure 7.1 shows the front-end of the pixel. We are interested in finding the root mean square (rms) noise value at the gate of the preamplifier input transistor. Artificially, the sensing process can be separated into three stages: the reset stage, the

sensing (integration) stage, and the readout stage. The noise sources during the sensing operation are listed below.

(1) During reset: The sources of noise come from thermal noise of the reset transistor.

(2) During integration: The sources of noise come from shot noise and flicker noise of the photocurrent and the dark current.

(3) During readout: The sources of noise come from thermal noise flicker noise of preamplifier.

Firstly, temporal noise in each stage is discussed in detail. Then, the sum of the noise in each stage is calculated, which gives the total temporal noise of a CMOS imager pixel.

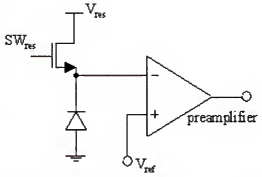



Figure 7.1 The diagram of the front-end of a pixel. Here, we are interested in finding the root mean square (rms) noise value at the gate of the preamplifier input transistor.

### 7.1.1 Reset Stage Noise

The photodiode voltage is first reset to the reset voltage before the integration of photogenerated charges. During reset, the gate of the reset transistor is biased to $V_{DD}$ for a few microseconds (a typical value for video applications, but it can be much longer for still imaging). The reset noise mainly comes from the reset transistor's channel thermal noise. Since the statistical uncertainties in the reset level and the ensuing signal level are highly correlated, effective signal processing circuits (for example correlated double

sampling) can efficiently reduce the reset noise. Figure 7.2 illustrates the model for reset noise analysis.

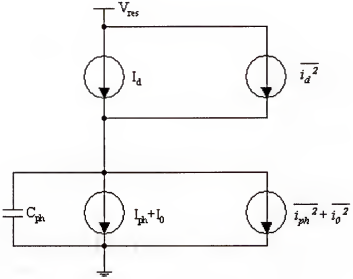



Figure 7.2 Noise model of reset stage.

In the figure $I_d$ denotes the reset transistor drain current. $\overline{i_d^2}$ denotes the thermal noise power associated with the reset transistor, and $\overline{i_{ph}^2}+\overline{i_0^2}$ denotes the shot noise power associated with the photocurrent and the dark current. Note that

$$I_d = I_{ph} + I_0 \tag{7.1}$$

is very small (typically less than 1nA). If the reset voltage is selected to be $V_{dd}$, then at the beginning of the reset the reset transistor is either operating in the saturation region or in the subthreshold region depending on the initial photodiode voltage at the end of integration. If the photodiode voltage is high, the reset transistor is in subthreshold. If the photodiode voltage is low enough, the reset transistor is in saturation at first and for a very short amount of time before it goes into subthreshold for the rest of reset. It is reported by Tian et al. [65] that the settling time is greater than 1 ms even for very high

photocurrents. This settling time is much larger than the typical reset time, which is typically in the few microseconds range, and this results in a KTC noise of at most $\frac{kT}{2C_{ph}}$ and image lag due to incomplete reset. In our design, the reset voltage is selected to be 3V. Therefore the reset transistor is always in above threshold range during reset and we do not need to worry about the image lag problem. Finally the reset transistor enters linear region and the channel resistance of the reset transistor can be estimated as

$$R_{ch} = \frac{1}{g_{ch}} = \frac{\partial V_{DS}}{\partial I_d} \approx \frac{1}{\mu_n C_{ox} \frac{W}{L}(V_{GS} - V_t)} \tag{7.2}$$

This value is estimated to be 794 Ω. Therefore the time constant for reset is about

$$R_{ch} C_{ph} = 794 \times 17.8 \times 10^{-15} \approx 14ps$$

which is very small and the reset processing reaches its steady state very quickly (within 1ns). So, the average noise power is simply the KTC noise power as shown below.

$$\overline{v_{res}^2} = \frac{kT}{C_{ph}} \tag{7.3}$$

At room temperature the KTC noise turns out to be

$$\overline{v_{res}} \approx 0.46mV$$

The reset noise (KTC noise) is independent of the photocurrent, dark current, and integration time.

### 7.1.2 Integration Stage Noise

In this stage, the sources of noise are shot noise and flicker noise of the photocurrent and the dark current. Shot noise is generated by fluctuations in the static (DC) current flowing through space charge regions, for example, the depletion region in

the photodiode pn junction. Flicker noise is associated with static (DC) current flow in both the resistive and depletion regions, caused by traps due to crystal defects and contaminants, which randomly capture and release carriers. The noise model for this stage is shown in Figure 7.3.

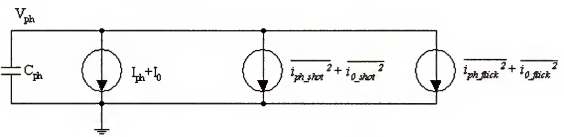



Figure 7.3 Noise model of integration stage.

The shot noise power due to the photocurrent and dark current is

$$\overline{i_{shot}^2} = 2q\left(I_{ph} + I_0\right)\cdot \Delta f \approx \frac{2q\left(I_{ph} + I_0\right)}{t_{\text{int}}} \tag{7.4}$$

where $t_{\text{int}}$ is the integration time of the pixel. The flicker noise can be expressed as

$$\overline{i_{ph,f}^2} = k_{f,ph}\frac{(I_{ph} + I_0)^a}{f^b}\Delta f \tag{7.5}$$

where the bandwidth $\Delta f$ is the inverse of integration time $t_{\text{int}}$. $K_{f,ph}$ is the noise constant for that device, $a$ is a constant with value between 0.5 and 2 and $b$ is a constant with value of about 1. These two noise sources are statistically independent.

### 7.1.3 Readout Stage Noise

The voltage on the photodiode is input to the preamplifier (see Figure 7.4) in the track mode of the comparator. In this stage the sources of noise are thermal noise and flicker noise of the preamplifier. There are two sources of the thermal noise in the image sensor. One comes from the conducting lines from pixels to the readout circuits. Since

digital signals are transferred in these lines, the thermal noise of these conducting lines can be ignored. The front-end conducting lines leading from the photodiode to the comparator are very short and here we also ignore their thermal noise. The other type of thermal noise is the MOSFET noise. It originates in the channel of the FETs, which dominates the other noise components of these transistors. This noise can be referred back to the gate of the FETs using the transconductance $g_m$ of the transistors, resulting in an equivalent gate voltage noise of the FETs.

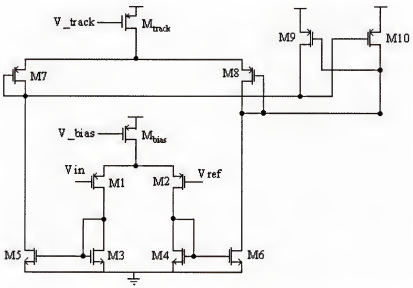



Figure 7.4 Preamplifier schematic.

The transistors in the signal path are involved in the noise analysis (see Figure 7.4). Due to the symmetry of the circuit structure, the signal path is half the circuit. Therefore, the input referred thermal noise power is

$$\overline{v_{eq,th}^2} = \frac{2}{g_{m1}^2}\left(\overline{i_{n1}^2} + \overline{i_{n3}^2} + \overline{i_{n5}^2} + \overline{i_{n7}^2} + \overline{i_{n9}^2}\right) \tag{7.6}$$

where $\overline{i_{nk}^2}$ ($k$=1,3,5,7,9) is the output noise power of transistor $M_k$ ($k$=1,3,5,7,9) and $g_{m1}$ is the transconductance of transistor $M_1$. Since the input stage of the preamplifier works in

subthreshold mode (very low bias current), the MOSFET transistor operates like bipolar transistor. With a simple but rough approximation, the dominant source of noise is estimated as

$$\overline{i_n^2} = 2qI_d \cdot \Delta f \tag{7.7}$$

where $I_d$ is the drain current and can be described as

$$I_d = I_s \frac{W}{L} e^{\frac{\kappa V_{GS} - (1-\kappa) V_{SB}}{V_t}} \left(1 - e^{-\frac{V_{DS}}{V_t}}\right) \tag{7.8}$$

here, $V_{GS}$ is the gate to source voltage, $V_{DS}$ is the drain to source voltage, $V_{SB}$ is the source to bulk voltage, $\kappa = \frac{C_{ox}}{C_{ox} + C_{depletion}}$ is the gate efficiency factor, $V_t$ is the thermal voltage and $I_s$ is a constant that depends on the transistor threshold voltage. The $g_{m1}$ in equation (7.5) can be estimated as

$$g_{m1} \approx \frac{I_{d1}}{V_t} \tag{7.9}$$

In our design, the input stage bias current is set to 700nA. The hand-calculated input referred thermal noise at room temperature turns out to be

$$\overline{v_{th,eq}^2} \approx 8.99 \times 10^{-15} V^2 / Hz$$

which is close to the SpectreS simulation result of $8.50 \times 10^{-15} V^2 / Hz$.

The input-referred flicker noise $\overline{i_{eq,f}^2}$ can be similarly estimated. Generally flicker noise dominates temporal noise for low frequency. However, it is negligible for the short integration times we are considering. P-type MOSFETs are used as input transistors. They provide less flicker noise than their NMOSFET counterparts. Some more details on flicker noise in solid state devices can be found in literature [66].

### 7.1.4 Total Temporal Noise

Since the noise sources are independent, the total temporal noise at the gate of the preamplifier input transistor is

$$\overline{v_{tot}^2} = \overline{v_{res}^2} + \overline{v_{th,eq}^2} + \frac{\overline{i_{shot}^2}}{C_{ph}^2} t_{\text{int}}^2 \tag{7.10}$$

### 7.1.5 Signal-To-Noise Ratio And Dynamic Range

To make the analysis tractable we make the following simplifying assumptions. First, as described early, we ignore flicker noise. In addition we model the input referred preamplifier noise as a band limited noise source with bandwidth B << $\omega_c$ (the cutoff frequency). In the consideration of signal-to-noise ration (SNR) and dynamic range (DR), dark current is also involved. Dark current is described by equation (1.13), here we rewrite it below

$$I_0 = Aq\left[n_i^2\left(\frac{D_p}{L_p N_D} + \frac{D_n}{L_n N_A}\right) + \frac{n_i W}{2\tau}\right] \tag{1.13}$$

where $n_i$ is very temperature-dependent. It doubles for every 11°C rise in temperature. The dark current density is ultimately limited by band-to-band recombination. Shrinking the pixel size and improving the quality of both the silicon wafers and the fabrication processes will lead to reduction in dark current and dark current noise. SNR can be described by

$$SNR = 10\log\frac{V_{ph}^2}{\overline{v_{tot}^2}} \tag{7.11}$$

$$= 10\log \frac{\dfrac{I_{ph}^2 t_{\text{int}}^2}{C_{ph}^2}}{\overline{v_{res}^2} + \overline{v_{th,eq}^2} + \dfrac{\overline{i_{shot}^2} t_{\text{int}}^2}{C_{ph}^2}} \tag{7.12}$$

$$= 10\log \frac{I_{ph}^2}{\dfrac{\left(\overline{v_{res}^2} + \overline{v_{th,eq}^2}\right) C_{ph}^2}{t_{\text{int}}^2} + \overline{i_{shot}^2}} \tag{7.13}$$

$$= 10\log \frac{I_{ph}^2}{\dfrac{\left(\overline{v_{res}^2} + \overline{v_{th,eq}^2}\right) C_{ph}^2}{t_{\text{int}}^2} + \dfrac{2q\left(I_{ph} + I_0\right)}{t_{\text{int}}}} \tag{7.14}$$

Equation (7.14) implies longer integration time results in higher SNR. So, we hope that a pixel should be integrated until it is saturated.

DR can be described by

$$DR = 20\log \frac{I_{ph,\max}}{\overline{i_{tot}}} \tag{7.15}$$

where $I_{ph,\max}$ is the maximum detectable photocurrent and $\overline{i_{tot}} = \dfrac{C_{ph}\overline{v_{tot}}}{t_{\text{int}}}$.

For conventional voltage-based CMOS imagers, the integration time for the whole pixel array is fixed. Using the parameters we used in our design ($V_{res}$ = 3V and $C_{ph}$ = 17.8fF) and assuming the integration time is 33ms, the maximum detectable photocurrent is about 1.62pA. Figure 7.5(a) and Figure 7.5(c) illustrate the SNR and contribution of noise sources for the voltage-based CMOS imager. It is clear that the SNR increases with $I_{ph}$ until the pixel enters the saturation state, where SNR equals about 52dB. Thermal noise contributes the least to the total noise, and reset noise (KTC noise) is the next. Both of them are constants. Shot noise is a function of signal and it increases

with $I_{ph}$. For SNR = 0dB, about 66dB dynamic range of this voltage-based CMOS imager can be achieved in this example.

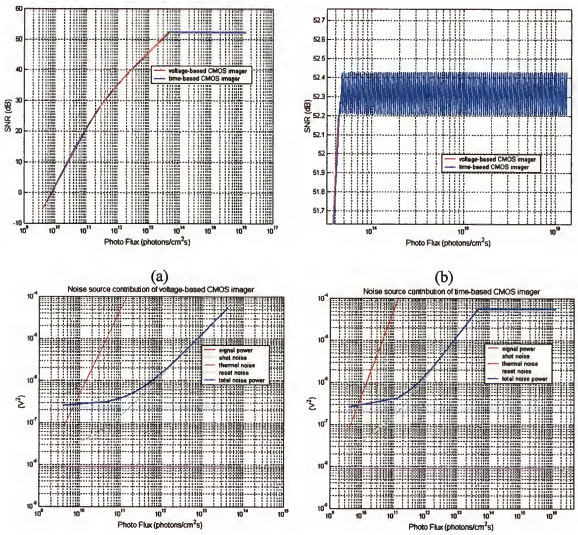



Figure 7.5 Noise analysis results. (a) The SNR of the voltage-based CMOS imager with integration time of 33ms and the SNR of the time-based CMOS imager with the longest sampling time of 33ms. (b) The SNR at high light intensity range. (c) The noise source contribution of the voltage-based CMOS imager. (d) The noise source contribution of the time-based CMOS imager.

For our time-based CMOS imager, the integration time is different for each pixel. Since it is a function of the pixel's incident light intensity. Short integration times are applied to high light intensity pixels, and long integration times are applied to low light

intensity pixels. Thus, wider dynamic range can be achieved in the time-based CMOS imager than the conventional voltage-based CMOS imager.

Figure 7.5(a) also illustrates the SNR of time-based CMOS imager for video applications. Figure 7.5(b) shows the details of the SNR for the high light intensity range. The comparison of the SNRs of voltage-based and time-based CMOS imagers shows that both of the two types of CMOS imagers have higher SNR in high light range but low SNR in low light range. The time-based CMOS imager has wider dynamic range than the voltage-based CMOS imager and the SNRs of the extended dynamic range are around 52.3dB. The dynamic range of the time-based CMOS imager is extended to about 120dB in this example. Figure 7.5(d) illustrates the noise source contribution of the time-based CMOS imager. It is similar to Figure 7.5(c).

## 7.2 Fixed Pattern Noise

Fixed pattern noise (FPN) is the spatial variation in pixel output values under uniform illumination due to device and interconnection parameter variations (mismatches) across the sensor. Therefore, it is also called non-uniformity. It is fixed for a given sensor, but varies from sensor to sensor. Assume $V_0$ is the nominal pixel output value at uniform illumination, and the output pixel values (excluding temporal noise) from the sensor are $V_{ij}$ ($i$ = 1,2, …, m, $j$ = 1,2,…,n), then the FPN is the set of values

$$\Delta V_{ij} = V_{ij} - V_0 \tag{7.16}$$

Unlike the temporal noise components we analyzed in section 7.1, FPN is a spatial noise. It affects SNR and causes more degradation in image quality at low illumination than high illumination. Generally, the sources of FPN of image sensors are spatial variation in photodetector geometry, column amplifiers (in voltage-based CMOS

imagers) and dark current. FPN for CCD image sensors appears random. The sources of FPN in a CCD imager are spatial variation in the photodetector geometry and the dark current. Neither CCDs nor the output amplifier that is shared by all pixels causes FPN. The sources of FPN in a voltage-based CMOS imager are pixel transistor and photodetector geometry variation, dark current variation and column amplifiers caused column FPN, which appears as "stripes" in displayed image. So, voltage-based CMOS images generally have higher FPN than CCD imagers. However, since time-based CMOS imagers output digital values, the FPN originates only in the front-end analog processing circuits. Comparison of SNR and FPN for wide dynamic CMOS image sensors can be found in literature [67]. The sources of FPN in our proposed time-based CMOS imager are the following.

(1) Variation in the photodiode geometry (e.g., area).

(2) Variations in reset transistor parameters (e.g., threshold voltage, it will affect charge injection).

(3) Variation in preamplifier parameters, (e.g., offset voltage).

(4) Variation in the dark current.

FPN is typically reported as the standard deviation of the spatial variation in pixel outputs under uniform illumination (excluding temporal noise) as a percentage of voltage swing. In the literature, FPN values of 0.1% to 4% have been reported [13-15,17-18,21]. Let $z_i$ denotes the random value of the parameter that is involved in FPN. It can be expressed as

$$z_i = \overline{z_i} + \Delta z_i \tag{7.17}$$

where $\overline{z_i}$ denotes the mean value (i.e. the nominal value) of $z_i$, and $\Delta z_i$ denotes the variation from the mean value. The mean of $\Delta z_i$ is 0 and standard deviation of $\Delta z_i$ is $\sigma_i$. Thus, the pixel output voltage value $V_0$ of the front-end circuit can be approximated by

$$V_0(\vec{z}) \approx \overline{V_0(\vec{z})} + \sum_i \frac{\partial \overline{V_0}}{\partial z_i} \cdot \Delta z_i \tag{7.18}$$

where $\vec{z}$ is the vector consisting of all the $z_i$'s, $\overline{V_0}$ is the nominal output voltage value, and $\frac{\partial \overline{V_0}}{\partial z_i}$ is the sensitivity of $\overline{V_0}$ with respect to the $i^{th}$ parameter evaluated at the nominal parameter value. Assuming that $\Delta z_i$'s are uncorrelated, FPN can be expressed as

$$FPN = \sqrt{\sum_i \left( \frac{\partial \overline{V_0}}{\partial z_i} \right)^2 \cdot \sigma_i^2} \tag{7.19}$$

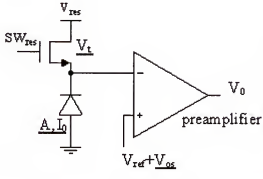



Figure 7.6 The parameters (highlighted with red underscore) involved in the FPN analysis.

Figure 7.6 shows the four parameters considered in the FPN analysis for our design. These parameters are highlighted with red underscore. $A$ denotes the photodiode area, $I_0$ denotes the dark current, $V_t$ denotes the threshold voltage of the reset transistor and $V_{os}$ denotes the offset voltage of the preamplifier. Since a dummy switch is used to reduce the charge injection from the reset transistor, we assume the effect of charge

injection is a linear function of $V_t$ as a first-order approximation. This is expressed as $k_{res}V_t$, where $k_{res}$ is a constant related with device parameters.

The output voltage in steady state can be expressed by

$$V_0 = A_v \cdot \left[ \frac{V_{res} - \left(J_{ph}A + I_0\right) \cdot t_{\text{int}}}{C_{ph}} + k_{res} \cdot V_t - V_{ref} - V_{os} \right] \tag{7.20}$$

where $A_v$ is the voltage gain of the preamplifier, $J_{ph}$ is photocurrent density, $t_{\text{int}}$ is the integration time and $C_{ph}$ is the capacitance at the cathode of the photodiode. The sensitivities of those involved parameters are estimated to be

$$\frac{\partial V_0}{\partial A} = \frac{A_v J_{ph} t_{\text{int}}}{C_{ph}} \tag{7.21}$$

$$\frac{\partial V_0}{\partial I_0} = \frac{A_v t_{\text{int}}}{C_{ph}} \tag{7.22}$$

$$\frac{\partial V_0}{\partial V_t} = A_v k_{res} \tag{7.23}$$

$$\frac{\partial V_0}{\partial V_{os}} = -A_v \tag{7.24}$$

So the FPN can be approximated as

$$FPN = \sqrt{\left(\frac{J_{ph} t_{\text{int}} \cdot \sigma_A}{C_{ph}}\right)^2 + \left(\frac{t_{\text{int}} \cdot \sigma_{I_0}}{C_{ph}}\right)^2 + \left(k_{res} \cdot \sigma_{V_t}\right)^2 + \sigma_{V_{os}}^2} \tag{7.25}$$

In the four terms of equation (7.25), the first and second ones are the contribution of photodiode geometry variation and dark current variation respectively. They are signal dependent. The third and fourth terms are the contribution of reset transistor threshold voltage variation and comparator offset voltage variation respectively. They are signal

independent. As an example, possible values of the $\sigma$'s of these four parameters are shown in Table 7.1 [68].

Figure 7.7(a) shows the contribution of the FPN sources for voltage-based CMOS imager with $t_{\text{int}}$ = 33ms, $C_{ph}$ = 17.8fF, and $k_{res}$ = 0.003. ($k_{res}$ is estimated based on SPICE simulation results. $k_{res} \approx \frac{V_{res} - V'_{res}}{V_t} = \frac{0.002V}{0.6586V} \approx 0.003$) The figure shows that the offset variation of the preamplifier/comparator dominates the FPN at low light intensity level and the variation of the photodiode geometry dominates the FPN at high light intensity level. Equation (7.21) shows that the sensitivity of photodiode geometry is proportional to photocurrent density. However, the SNR is large enough at high light intensity level but is small at low light intensity level. Therefore, FPN at low light intensity is more important. Thus, to reduce the effect of FPN, offset voltage variation, which is the major contribution of FPN at low light intensity range, should be cancelled or calibrated. This will be discussed in section 7.4. The maximum FPN is about 10mV. If the signal swing is 3V, the FPN is 0.33% of signal swing.

Table 7.1 The standard deviation of parameters

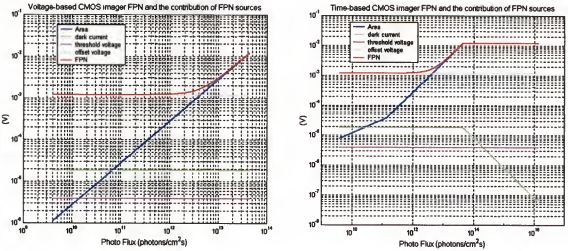
| Parameter | σ |
|---|---|
| A | $0.4\% \cdot \overline{A}$ |
| $I_0$ | $2\% \cdot \overline{I_0}$ |
| $V_t$ | $0.2\% \cdot \overline{V_t}$ |
| $V_{os}$ | 1.2mV* |

*: This is based on SPICE (Bisim 3.3 version, level 49) simulation result.

Figure 7.7(b) illustrates the FPN and the contribution of FPN sources of the time-based CMOS imager with the implementation of Weber's law. Similarly, offset voltage variation of the preamplifier/comparator still dominates the FPN at low light intensity range and photodiode geometry variation dominates the FPN in high light intensity range.

However, the contribution of dark current variation is not constant due to different sampling time for different photocurrent. The maximum FPN is also about 10mV, similar as that of the voltage-based CMOS imager. Therefore, the FPN is 0.33% of signal swing.

Also notice that the offset variation induced FPN dominates when integration time is not long. When integration is long, for example 2.5s, the dark current induced FPN will dominates. As we will see in next Chapter, the dark current induced FPN corresponds to 3.3% of signal swing, which is much greater than offset induced FPN of 0.22% of signal swing.




(a) (b)

Figure 7.7 FPN and the contribution of FPN sources. (a) For voltage-based CMOS imager. (b) For the time-based CMOS imager with the implementation of Weber's law.

### 7.3 Defects

Image quality is one of the important specifications of an image sensor. This means that the sensor's image should be flawless. Unfortunately, technology is not perfect. Due to imperfections in the processing, a small number of pixels in an array may be defective (hard faults) or operate out of the specification limits. Such pixels appear as white or black points in the image. Typically, the white spots are caused by impurities in

the silicon. These yield locally a very high dark current, which cannot be distinguished from the photocurrent. For a human observer, these tend to be much more annoying than other image imperfections such as temporal noise and a mild FPN. Normally, such a pixel fault can be isolated and the error can be corrected by interpolating the intensity information from neighboring pixels. Dierickx et al. proposed an algorithm for detection and correction of missing pixels [69, 70] that interested reader can refer to and we will not discuss the detail of defects correction.

### 7.4 Issues Of Off-chip Offset Calibration And On-chip Offset Cancellation

Since offset voltage is the most important source of FPN, it should be canceled on-chip or calibrated off-chip. Correlated double sampling (CDS) is a commonly used method in analog circuits to reduce the effect of offset voltage on chip. It is also suitable in CMOS imagers to reduce the FPN and reset noise. In CDS, the output is sampled twice: once with the signal present and the second time immediately after reset. Then, the offset voltage and reset noise are eliminated by taking the difference of the two samples of the output. However, CDS can only reduce the signal independent parts of the FPN. The signal dependent part of the FPN cannot be reduced using CDS. In DPS pixel design, CDS circuits have to be integrated inside each pixel and must include a high gain amplifier to realize an auto-zeroing procedure. Since a low-gain preamplifier is implemented to achieve wide frequency response, an in-pixel CDS is not applicable for our case. Here, we propose an off-chip offset calibration method.

Figure 7.8 illustrates the principle of one approach to realize the offset calibration. The negative input of the comparator is set to the properly selected reset voltage $V_{res}$ (e.g., $V_{res} = \frac{V_{dd}}{2}$). To do this, the reset transistor has to be in the "on" state during

calibration. The positive input of the comparator is the reference voltage plus the offset voltage ($V_{ref}+V_{os}$). The reference voltage is selected to be in the range of $[V_{res}-\psi, V_{res}+\psi]$. The reference voltage is swept linearly in its range with a step size of $\Delta V_{ref}$. At each sweeping step, the array is latched and the status of each pixel is read out such that the offsets of them are identified and stored.

The output of the comparator toggles when

$$V_{res}=V_{ref}+V_{os} \tag{7.26}$$

so

$$V_{os}=V_{res}-V_{ref} \tag{7.27}$$

In calibration, the offset voltage caused time variation can be calculated as

$$t_{off}=-\frac{V_{os}\cdot C_{ph}}{I_{ph}} \tag{7.28}$$

Since

$$I_{ph}=\frac{C_{ph}\left(V_{res}-V_{ref}\right)}{t_{\text{int}}} \tag{7.29}$$

We can obtain

$$t_{off}=\frac{V_{os}}{V_{ref}-V_{res}}\cdot t_{\text{int}} \tag{7.30}$$

Here, $V_{ref}$ and $t_{\text{int}}$ are known parameters that can be found in the $V_{ref}$ table and the timing table. $V_{res}$ and $V_{os}$ are also known. So, $t_{off}$ can be calculated based on equation (7.30). The calibrated time will be $(t_{\text{int}}+t_{off})$. Remember the variation of $V_{os}$ matters, not the absolute value of $V_{os}$.

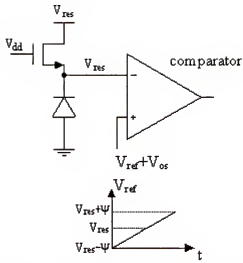



Figure 7.8 The principle of off-chip offset calibration.

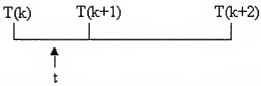



Figure 7.9 Index assignment: (N-k-1) is assigned to pixels with integration time t.

The difficulty of off-chip calibration for the time-based CMOS imager is that offset calibration is not accurate due to the quantization of the photocurrent. This may result in over-calibration and/or under-calibration. Also, off-chip calibration does not reduce KTC noise as CDS does in conventional APS architecture. Assume the expected time $t$ is located between sampling time $T$(k) and $T$(k+1) as shown in Figure 7.9, the time-based sampling will assign the index (N-k-1) to the sample. We will obtain the following calibration results:

- If $T(k-n-1) > T(k+1) + t_{off} \geq T(k-n)$, where $t_{off} < 0$ and $n = 0,1,\ldots$, then index (N-k+n) should be assigned to the pixel.
- If $T(k+n+1) > t_{off} + T(k) \geq T(k+n)$, where $t_{off} > 0$ and $n = 1,2,\ldots$, then index (N-k-n-1) should be assigned to the pixel.

- Otherwise, there could exist one index error for the pixels. We cannot accurately calibrate the offset voltage because we actually do not have the analog voltage output.

Note that even for the first two calibrations, there still exist errors because the $t_{off}$ is a function of photocurrent (see equation (7.28)). The calibration just tries to make the effect of offset smaller. For accurate calibration, on-chip offset cancellation need to be involved which directly deal with the analog signal.

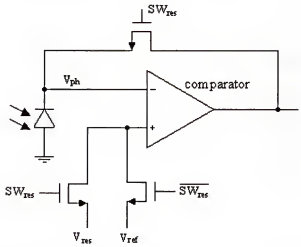



Figure 7.10 Revised DPS pixel with the on-chip offset cancellation.

To realize on-chip offset cancellation for our time-based CMOS imager, a revised CDS circuit can be designed as shown in Figure 7.10. In the figure, the comparator has a high preamplifier gain of more than 40dB (compared with 12dB of previous design) such that CDS processing can be correctly processed. This comparator bandwidth reduces to 4% of the previous design. When $SW_{res}$ is $V_{dd}$, the preamplifier of the comparator works as a voltage follower and the photodiode voltage $V_{ph}$ is reset to $V_{res} + V_{os}$. When $SW_{res}$ switches to ground, the reference voltage $V_{ref}$ is connected to the positive input of the comparator and the system enters integration mode. Since $V_{os}$ appears at both the positive

and negative inputs of the comparator, offset voltage is cancelled very well. This on-chip offset cancellation technology sacrifices the bandwidth of the comparator and may cause a large system error as shown in Figure 5.4.

CHAPTER 8
SIMULATION AND EXPERIMENTAL RESULTS

In this chapter, we present the Cadence SpectreS simulation and experimental results of the circuits and system. Single pixel simulation and experimental results are presented first. Then, the simulation and experimental results of the readout circuits with the configurable shift register architecture are presented. Finally measurements from a 32×32 pixel array are presented.

## 8.1 Single Pixel Simulation And Experimental Results

Figure 8.1 shows the diagram of a pixel. It consists of a photodiode, a reset transistor, a comparator with disabling circuit and a row select transistor. Simulation and experimental results of the photodiode, comparator and a single pixel are presented in the following subsections.

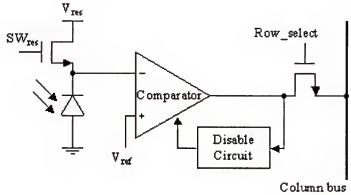



Figure 8.1 Diagram of a single pixel.

To test the performance of the photodiode, a square wave ranging from 0V to 5V is applied to the gate of the reset transistor that acts as a switch. This square wave is the control signal $SW_{res}$ in Figure 8.1. When $SW_{res}$ is 5V, the switch is closed and the photodiode is reset. When $SW_{res}$ is 0V, the switch is opened. The photogenerated current discharges the photodiode cathode capacitance, approximately following the relation of $C_{ph}\frac{dV_{ph}}{dt} = I_{ph}$. Figure 8.2 shows the testing result under the luminance of 2400lux. The reset voltage $V_{res}$ = 3V and the period of the $SW_{res}$ is 0.36ms. The photodiode voltage drops linearly versus time as expected. Note that due to the lack of specialized experimental equipment, the accuracy of the measured results may be further improved.

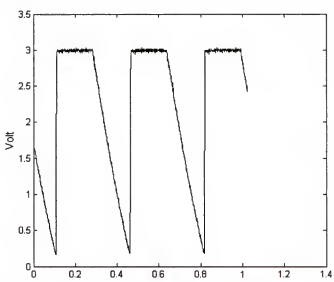



Figure 8.2 Photodiode test result: photodiode voltage versus time.

Dark current is measured for a single pixel. The chip is tested in dark environment and reset to 3V. Then the photodiode is discharged by dark current. Reference voltage is set to be a constant value. The times for photodiode voltage to drop to reference voltage are measured. This experiment was done 20 times. The mean and variance of the time is

calculated. The variance is caused by temporal noise and readout error. The experiment was done under three different bias voltages cases and the results are shown in Table 8.1.

Table 8.1 Experimental results for dark current testing (room temperature).

| $V_{bia}$ | $I_0$ | Temporal noise (% of signal) | $SNR_{simu}$ | $SNR_{test}$ |
|---|---|---|---|---|
| 4.8V | 3.16fA | 0.18% | 52.0dB | 54.8dB |
| 4.2V | 3.39fA | 0.20% | 52.0dB | 54.0dB |
| 3.8V | 7.60fA | 0.23% | 52.7dB | 50.0dB |

Note that the dark current measured here includes the leakage current in the photodiode pn junction and the leakage current introduced by substrate carriers. The measured total dark current value is higher than that shown in the literature [50]. One reason is that the large pixel that has much more complex analog and digital circuit injects more carriers into substrate. These carriers are absorbed by the pn junction of the photodiode and form a current in parallel with the dark current. Therefore, when the bias current increases, there exists more substrate-induced leakage current. Another reason comes from the current due to voltage bounce on the supply/return rails [71]. Experimental results show that as bias voltage decreases from 4.8V to 4.2V and to 3.8V, the measured leakage current increases from 3.16fA to 3.39fA and to 7.86fA. To reduce this effect, a small pixel with simple circuit design will be helpful. Another method is to use an $n^+$ guard ring around the n-well/p-sub photodiode to absorb the substrate carriers. This can dramatically reduce the effect of substrate carriers.

The temporal noise comes from the KTC noise, shot noise of dark current, thermal noise of comparator and power supply voltage variation. The experiment was done in dark environment to avoid room light variation. The measured SNR are 54.8dB, 54.0dB and 50.0dB respectively. To compare them with theoretic results, equation (7.14) can be revised to the following equation:

$$SNR = 10\log\frac{I_0^2}{\frac{\left(\overline{v_{res}^2} + \overline{v_{th,eq}^2}\right)C_{ph}^2}{t_{\text{int}}^2} + \frac{2qI_0}{t_{\text{int}}}} \tag{8.1}$$

Substituting the noise source values and the integration time used in the experiments, we can obtain the simulated SNRs that are listed in Table 8.1. The simulation results and experimental results are close to each other. The difference mainly comes from the power supply variation and substrate noise that are not included in the theoretic model in Chapter 7.

To measure the dynamic range of a single pixel, a pixel is exposed to various light intensities. The reset voltage is set to 3V. The light intensities in the environment are measured using digital light meter LX-102. The reference voltage is set to be low in the high and middle light intensity range but high in the low light intensity range. The integration times needed for the comparator to toggle are measured using the digital oscilloscope Tektronix 2430A. To demonstrate the concept of extending dynamic range based on two-degrees of freedom and keep a relatively high SNR in the measurement, the sampling times and reference voltages, the maximal integration time is set to 300 ms in the experiment. The shortest sampling time is measured to be 8.1 μs. After the experiment, the light intensities are reconstructed from the timing and reference voltage information. The sampling time, reference voltage and reconstructed light intensity versus the measured light intensity are shown in Figure 8.3. The regression analysis shows that the adjusted $R^2$ of reconstructed light intensity to measured light intensity is 0.99. The measured dynamic range is 115dB, to which sampling time contributes 88dB and reference voltage contributes 27dB.

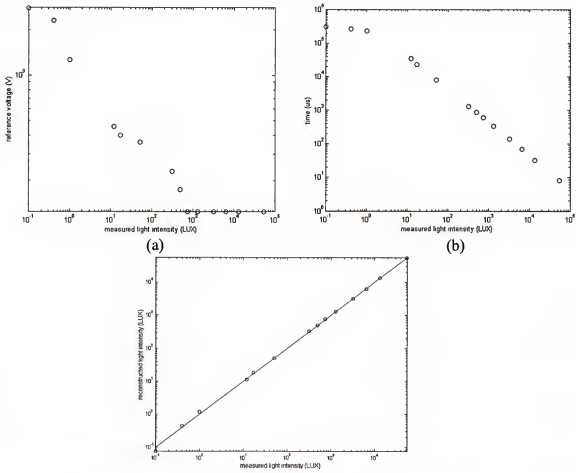



Figure 8.3 Single pixel test results: (a) time versus measured light intensity. (b) reference voltage versus measured light intensity. (c) reconstructed light intensity versus measured light intensity (circles: sampling points; solid line: regression line).

The sensitivity [72] of the pixel is decided by

$$s = \frac{\Delta V}{\text{incident_light_intensity} \cdot \Delta T} \tag{8.2}$$

From the experiment, the sensitivity of a pixel can be found to be about 6.4V/lux·s.

### 8.2 Readout Circuit Simulation And Experimental Results

A 2×10 pixel array absolute address coding (AAC) readout architecture was simulated using SpectreS and also fabricated using AMI 0.5um processing and tested.

The external clock signal is used directly as the clock of the column shift registers (CSR) or $CSR_clk = clk$. The clock of the row shift registers (RSR) is generated by

$$RSR_clk = \overline{A_0 \cdot end_mark \cdot \overline{sample} \cdot CSR_clk}$$

The bit-line values of the CSR are controlled externally and have the logic value of "1" or "0". The addresses of $A_0$ are set to {0,1,0,1,0,1,0,1,0,1}. This circuit is used to test the logic function of the AAC architecture. Figure 8.4 shows the simulation results with all the bit-line values are set to logic "1". When the sample single is high, the circuit is ready for data readout: The load_e becomes high to load the column addresses to the column shift registers; the RSR_clk becomes clk such that the row shift registers are loaded. After the sample signal becomes low, the addresses start to be read out. At the end of one row of data, the end_mark (here it is logic high) appears to indicate the end of the row and trigger the RSR_clk and load_e signals to become high such that the column addresses are re-loaded. After the end_mark is read out, the RSR_clk becomes low and its up-down edge cause the row shift registers to shift such that the next row is selected and start to read out the addresses of it. The RSR_out here is the end-frame-bit. Here it is also used to indicate the function of the row shift registers. Clearly it has two RSR_clk periods low state and one RSR_clk period high state. This comes from the two rows structure of the circuit.

Figure 8.5 shows the experimental results of the circuit. A slow frequency clock was used in the experiment such that the measured frequencies are within the digital oscilloscope's bandwidth. In this experiment shown in the figure, all the column lines were set logic high and we can see that there are 10 clocks between RSR_clk, which is what we expected. The addresses of $A_0$ are sent out in series and the end_mark appears at

the end of the row address sequence and is read at the clock period covered by RSR_clk. After two RSR_clk pulses, the RSR_out changes from logic low to logic high and stays in the logic high state for one RSR_clk period and then becomes low again. This exactly corresponds to the simulation results. The load_e signal behaves as RSR_clk. Since the address data are correct for each row, the "load_enable" function is correctly accomplished. This indicates that at the end of one row address readout, the column addresses are re-loaded in the column shift registers. This is consistent with the simulation results.

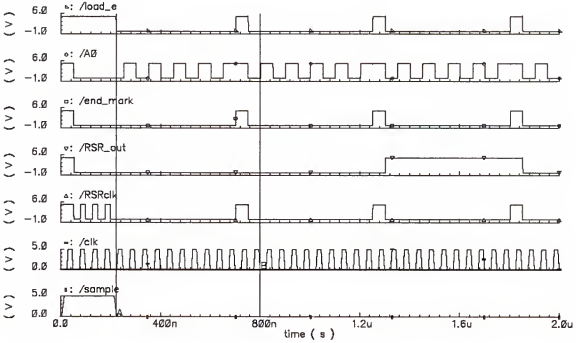



Figure 8.4 Simulation results of the timing signal of AAC readout architecture.

The layout of the test chip is shown in Figure 8.6. It is layouted using AMI 0.5um technology. The circuitry at the top is a single pixel testing circuitry. The photodiode voltage is read out through a follower-connected opamp. The testing point signals are also readout through follower-connected opamps. The circuitry at the bottom of the chip

is the 2×10 pixel array absolute address coding readout circuitry. The pixel outputs are controlled externally. There are totally 40 pads in the testing chip.

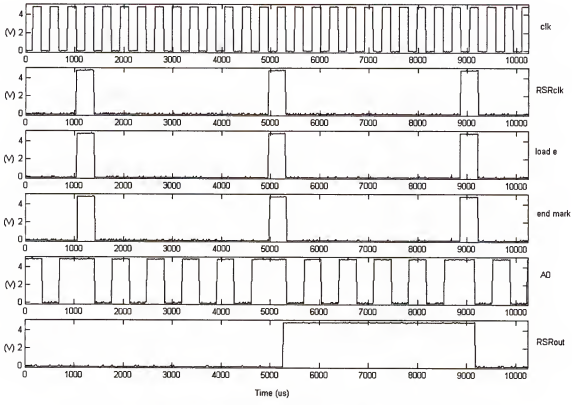



Figure 8.5 Measured timing signal of AAC readout architecture.

Figure 8.6 Layout of the test chip.

### 8.3 Image Sensor Array Simulation And Experimental Results

A 2×2 pixel array circuit was simulated in Cadence SpectreS. Sixteen D flip-flops (in column shift register) are included in the simulation. Constant current sources are used as the photocurrents to discharge the photodiode's parasitic capacitor. The control logic is same as that used in the chip of a 32×32 pixel array. An extra row is added to the bottom of the pixel array with a single grounded NMOS transistor in each pixel to guarantee the last row of the sensor array can undergo the disabling processing. The gates of these NMOS transistors are connected to the row_select and the drains of these transistors are connected to the column lines. When this extra row is selected, the only data that will be read out is the "end_mark" and meanwhile, the last row of the sensor array is under the disabling processing. At the next RSR_clk cycle, the "end_frame_bit" becomes logic "1" and the array enters track mode. If the extra row is not included, the array enters track mode with the last row undergoes disabling processing. This will result in the pixels in the last row are disabled in the track mode of first sampling before they are read off chip and therefore, the last row becomes a bad row. In the simulation design, we expect that pixels located at (1,1), (1,2) and (2,1) should be read out at the first sampling time and pixel located at (2,2) should be read out at the second sampling time. After each sampling, the fired pixels should be disabled and should not be read out again in the following sampling times. Figure 8.7 illustrates the simulation results. They are consistent with what we have expected.

Figure 8.8 shows the schematic of a 32×32 pixel array CMOS imager circuit. It has three major parts: pixel array (within the square box), logic control (within the polygon box) and row and column decoders (the rest).

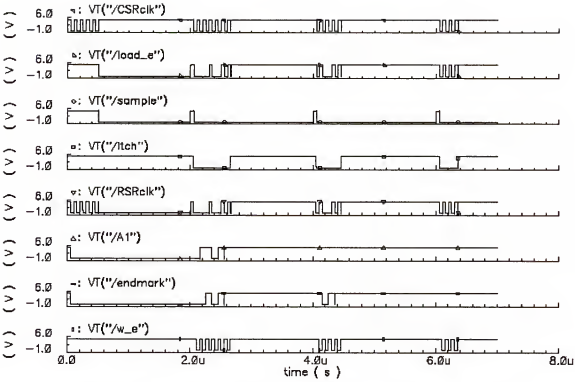



Figure 8.7 The simulation results of the 2×2 pixels array testing circuit.

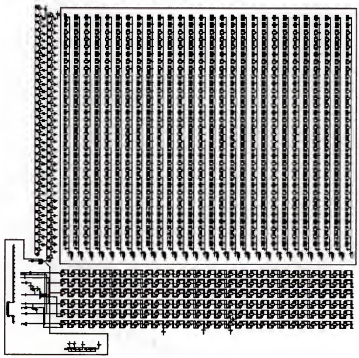
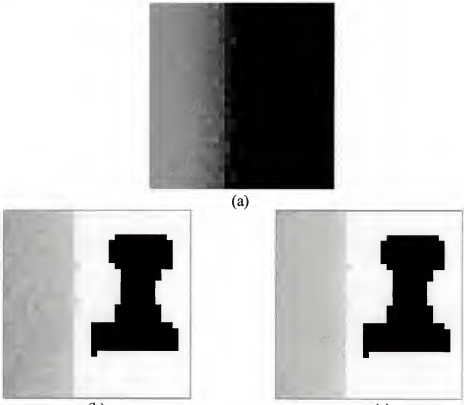


Figure 8.8 Schematic diagram of the 32×32 pixel array CMOS imager.

Several measured results are shown in Figure 8.9. The index image in Figure 8.9(a) shows the experiment result of the imager array under room light environment with half of the array being covered. The dynamic range of this image is 85dB. A letter "I" taped on a long fluorescence lamp was captured using a 4mm lens mounted on the imager and the index image is shown in Figure 8.9(b). The dynamic range of this image is 104dB. The post-processing image where the variation of dark current is subtracted from Figure 8.9(b) is shown in Figure 8.9(c). It looks smoother than the raw image.

FPN is evident in these experimental results. The FPN comes mainly from the offset voltage variation of the comparators. In the current design, we do not include off-chip offset calibration function and we do not implement on-chip offset cancellation design either. This results in the evident FPN. On-chip offset cancellation design should be taken into consideration in future designs. To measure the FPN, uniform light sources should be used. Since we do not have this kind of device in the lab, we cannot measure the FPN characteristics under uniform light source. However, FPN characteristics in dark environment can be measured and this will give us a quantitative image of FPN of CMOS image sensors. Figure 8.10 shows the space variation of the dark currents. In the experiment, the pixel comparator bias voltage is set to 3.8V. The dark image was captured 20 times. Then average of the images was calculated such that temporal noise effect can be reduced. The mean value of the dark current is 6.35fA and the standard deviation is 0.75fA. So this FPN under dark is about 3.3% of signal swing. The standard deviation of offset voltage is 6.65mV. It contributes FPN of 0.22% signal swing, which is much less than 3.3%. This is consistent with what we have mentioned in Chapter 7 that

when integration is long, dark current induced FPN dominates. In summary, the characteristics of the CMOS imager are listed in Table 8.2.



(a) (b) (c)

Figure 8.9 Index image results of the 32×32 pixel array CMOS imager. (a) The imager array under room light environment with half of the array being covered. Dynamic range is 85dB. (b) A letter "I" taped on a long fluorescence lamp and was captured with a 4mm lens mounted on the imager array. Dynamic range is 104dB. (c) The post-processing image where the variation of dark current is subtracted from Figure 8.9(b).

In the AMI 0.5um processing, there are three metal layers and one polysilicon layer. Generally, in the layout design, the metal_1 is placed in horizontal direction and metal_2 is placed in vertical direction. The metal_3 layer is used to cover the layout area with opens to the photodiode area. The layout of a single pixel, a column shift register unit, a row shift register unit and the 32×32 pixel array CMOS imager chip layout are shown in Figure 8.11. The area of the chip is 2940um×2940um. There are 40 MOSIS-provided pads.

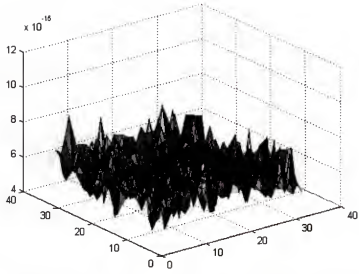



Figure 8.10 Space variation of dark current in the 32×32 pixel array.

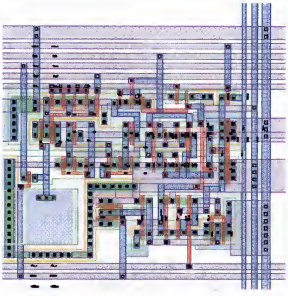
Table 8.2 Characteristics of the CMOS imager

| | |
|---|---|
| Process | 0.5μm 5V CMOS, single poly, triple metal |
| Pixel pitch | 60μm ×60μm |
| Pixel fill factor | 2.8% |
| Saturation voltage | 3V |
| Sensitivity | 6.4V/lux·s |
| Space resolution | 32×32 |
| Dynamic range | 115dB |
| Temporal noise | 0.20% |
| FPN | 3.3% (dark) |
| Power consumption | 37.8mW |

## 8.4 Discussion

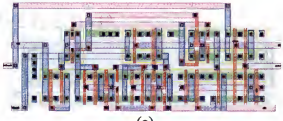
As the photodiode is placed close to the circuit, "cross talk" happens during the process. The electrons in the substrate may be absorbed by the photodiode. This will cause an increase of dark current. One way to minimize the effect of "cross talk" is to add $n^+$ guard ring around the photodiode.

The power supply of mixed-signal circuit can cause trouble to analog circuit part due to the power supply variation caused by the digital circuit part. Separating power supply can help avoid this problem.

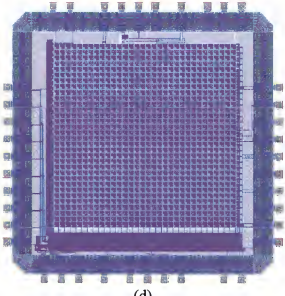


(a)

(b)



(c)



(d)

Figure 8.11 Layout of units. (a) A single pixel. (b) A row shift register unit. (c) A column shift register unit. (d) CMOS imager chip layout.

The fill factor of the current design is 2.8%. The wide pitch between pixels causes bad space resolution of the captured images. A more dedicated design could reduce the transistor numbers and thus help increase the fill factor. Design with smaller size processing, for example, 0.18um technology, can also help increase fill factor. Some special technology such as thin film on ASIC (TFA) provides different layers for sensing and process and it provides 100% fill factor.

# CHAPTER 9
# CONCLUSION AND FUTURE DIRECTION

## 9.1 Conclusion

A time-based synchronous readout CMOS imager is proposed in the dissertation. It provides two-degrees of freedom, i.e. the sampling time and reference voltage, to realize ultra-wide dynamic range imaging. We also proposed optimal sampling strategies to maximize the signal-to-noise ratio (SNR). The sampling interval is constrained by the readout speed and must be longer than the worst readout time of a frame (i.e. the minimum sampling interval). This constraint is named time domain minimum contrast constraint. To maximizing the SNR, full well charge should be discharged and this is named full well charge (FWC) criterion. The sampling strategies decided by these two criteria achieve maximum SNR and must satisfy the time domain minimum contrast requirement. In the applications without sampling time limitation such as still imaging, constant reference voltages of zero should be applied and the sampling times are decided by the quantization of the light intensities and the minimum contrast law. For the cases where time domain minimum contrast law is violated, reference voltage has to be increased until the time domain minimum contrast constraint is satisfied. This results in a partial use of the whole well charge. This reference voltage adjustment usually occurs in the high light intensity range. The bending-up point in high light intensity range is named $\alpha$. In applications with limited sampling time such as video applications, reference voltages will bend up towards the reset voltage in the low light intensity range. The

sampling time interval in this reference voltage bending-up range should be set to the minimum sampling interval. The bending-up point in low light intensity range is named $\beta$. The application of linear quantization of light intensity and the application of Weber's law are examined in detail in this dissertation.

Temporal noise and fixed pattern noise (FPN) are analyzed in detail. The noise sources come from the front-end analog circuit part. The sources of KTC noise, shot noise and thermal noise are analyzed from three stages: reset stage, integration stage and readout stage. SNR and dynamic range are analyzed for conventional voltage-based CMOS imagers and the time-based CMOS imager. It turns out that the time-based CMOS imager has wider dynamic range than conventional voltage-based CMOS imagers and keeps almost the same SNR performance in low and medium light intensity ranges and better SNR in high light intensity range. FPN are analyzed based on the statistical model we used. It turns out that the offset voltage variation of the comparators is the dominant source of FPN and we need to calibrate the output using an off-chip offset calibration method or an on-chip offset cancellation method.

Different high-speed readout schemes are also proposed and the relative and absolute address coding (R&AAC) technique turns out to be the one that has the fastest readout speed with a little bit more gate delay and reconstruction complexity. Only the fired pixels are read out and each pixel is read out just once. For N multiple samples, the R&AAC readout scheme increases the readout speed by N times than conventional bit frame readout scheme when no multiplex processing is involved.

Circuit design is discussed in detail and the system-level architecture is presented. The whole system is built on a 2.94×2.94 $mm^2$ silicon using AMI 0.5um technology. A test chip is also built to test the performance of different part of the system, i.e. the pixel, the comparator and the readout architecture. Cadence Spectre simulation results match the measurements well. However, the FPN is evident in the captured images due the lack of offset calibration circuit.

In summary, the time-based CMOS imager has the following advantages:

1) Ultra-wide dynamic range imaging
2) High speed readout strategy: bit plane compression readout scheme
3) Pixel sensor with direct 1-bit digital output: no need for ADC
4) Straightforward image reconstruction.

The major contributions of this research work are:

1) The proposal of two degrees of freedom time-based sampling technique
2) The analysis of the optimal sampling strategy for time-based multiple sampling technique
3) The development and analysis of the bit plane compression readout schemes
4) The noise analysis of the time-based CMOS imager
5) The successful design and test of the fabricated system

## 9.2 Future Directions

There are two important issues remaining to investigate. First, to achieve more accurate output, on-chip offset cancellation technique should be involved. One solution is shown in section 6.4 of Chapter 6. This will result in a longer comparator settling time

due to the finite unity gain frequency of a comparator. Advanced technology such as 0.18um technology, helps increase the unity gain frequency.

Second, to make this design more practicable, adaptively sensing light luminance should be involved. We do not need to fix the sensing range to a particular range like from $I_0$ to $10^6 I_0$. The first fired pixels indicate the brightest light intensity in the scene and the quantization is applied to the range from $I_0$ to the corresponding maximum photocurrent. This will increase the complexity of the system. A sampling start-up circuit needs to sense the first firing time with reference voltage of 0. A more involved method would continue to sample only when significant numbers of pixels fire. Such adaptive strategies will make the system more accurate and more practicable.

APPENDIX A
THE MATLAB CODE FOR THE DEMO OF IMAGE "LENA"

```
% This program is an example to sample an 8-bit
% image with nonlinear 6 bits representation.
% delta_v is calculated based on the 64 selected Iph and
% sampling time % (in this    example
% sampling time is linearly increased).
% The transform from nonlinear representation to linear representation % is examined.

% Qiang Luo
% May 28, 2001

clear
close all
N=64;
Iph=zeros(N,1);

% ---------------- set Iph -------------
% Iph<255 & Iph>=211
step=11;
for k=1:5
   Iph(k)=255-(k-1)*step;
end
last=5;

% Iph<211 & Iph>=92
step=4;
for k=1:27
   Iph(last+k)=200-k*step;
end
last=32;

% Iph<92 & Iph>=11
step=3;
for k=1:27
   Iph(last+k)=92-k*step;
end
last=59;
```

```
% Iph<11
Iph(60)=8;
Iph(61)=5;
Iph(62)=2;
Iph(63)=1;
Iph(64)=0;

% ------------------- Iph set finished ----------------

% calculate the delta_v
% assume the time has a constant step of 1. Thus we have time instance from 1 to 64
% for the 64 samplings
t=1:N;
t=t';
% assume a normalized capacitance, thus the C=1
C=1;
delta_v=zeros(N,1);
delta_v=Iph.*t/C;
f=1./t;

% verify the reconstruction
load lenademo.mat
yy=zeros(256,256);
for m=1:256
   for n=1:256
      yy(m,n)=f(N-y(m,n))*C*delta_v(N-y(m,n));
   end
end

% reconstructed image has exactly the same value as original sampling
% value
z=x-yy;
min(min(z))
max(max(z))

load lena.mat
figure
image(girl)
axis square
axis off
colormap(gray(256));
title('original image')

figure
subplot(1,2,1)
image(y)
```

```
axis square
axis off
colormap(gray(64));
title('sampled image')

subplot(1,2,2)
image(yy);
axis square
axis off
colormap(gray(64));
title('reconstructed image')

% The test for high dynamic range
for k=0:8:(256-64)
   z=yy-k;
   subplot(1,2,2)
   image(z);
   axis square
   axis off
   colormap(gray(64));
   title('reconstructed image')

   pause
end

for k=(256-64-8):-8:0
   z=yy-k;
   subplot(1,2,2)
   image(z);
   axis square
   axis off
   colormap(gray(64));
   title('reconstructed image')

   pause
end
```

APPENDIX B

THE MATLAB CODE FOR THE DEMO OF WIDE DYNAMIC RANGE IMAGES

```
% This program is a demo to display a wide dynamic range image.
% 8 bits record is used.
%
% The transform from nonlinear representation to linear representation % is examined.

% Qiang Luo
% July 17, 2001

% ------------- examine the relations among Iph, Vref, and tint ------
clear
load Vref_8bits.mat
Vres=2.5;                    % units: volt
delta_v=Vres-Vref;

% ------------ verify the reconstruction ----------------------
% y is the nonlinear reconstructed image. It consists of sampling index
% These can be used to search the sampling time table and reference
% voltage table to reconstruct an image - linearized reconstructed image.

load suntreedemo.mat
% x is original image, y is sampled image
yy=zeros(size(y));  % yy is linearized reconstructed image: Iph
[M,N]=size(yy);
C=1e-5;          % C is estimated based on the Iph_ratio table
fint=1./tint;

% here max(max(yy))=1, in other words, yy is normalized Iph
% sampling table. For display the matrix should multiply some value
% e.g. 256
for m=1:M
   for n=1:N
      yy(m,n)=fint(256-y(m,n))*delta_v(256-y(m,n))*C*10^6*162.5;        % 10^6: MHz
   end
end

figure
subplot(1,2,1)
```

```
y=1.497.*y;
image(y)
axis off
axis equal
colormap(gray(256));
title('sampled image')

subplot(1,2,2)
yy=360*256.*yy;
image(yy);
axis off
axis equal
title('reconstructed image')

% The test for high dynamic range : invisible parts
for k=64:64:3136  % this is for suntreedemo.mat
   z=yy-k;
   subplot(1,2,2)
   image(z);
   axis off
   axis equal
   title('reconstructed image')
   pause

end

for k=3136-64:-64:64
   z=yy-k;
   subplot(1,2,2)
   image(z);
   axis off
   axis equal
   title('reconstructed image')
   pause

end
```

## REFERENCES


[1] Q. Luo, Z. Chen, J. Harris, S. Clynes and M. Erwin. A Two-Degrees of Freedom Time Domain Sampling Technique for CMOS Imager. Patent filed in Texas Instruments Inc., August, 2001.

[2] Q. Luo, Z. Chen, J. Harris, S. Clynes and M. Erwin. Adaptive Relative and Absolute Address Coding CMOS Imager Technique and System Architecture. Patent filed in Texas Instrument Inc., August, 2001.

[3] E. Fossum. CMOS Image Sensors: Electtronic Camera-On-A-Chip. IEEE Trans. On Electron Devices, vol. 44, no. 10, pp. 1689-1698, 1997.

[4] G. Weckler. Operation of p-n Junction Photodetectors in a Photon Flux Integrating Mode. IEEE Journal of Solid-State Circuits, col. Sc-2, no. 3, pp. 65-73, 1967.

[5] D. Scheffer, B. Dierickx, and G. Meynants. Random Addressable 2048×2048 Active Pixel Image Sensor. IEEE Trans. On Electron Devices, vol. 44, no. 10, pp. 1716-1720, 1997.

[6] E. Stevanovic, M. Hillebrand, B. Hosticka, and A. Teuner. A CMOS Imager Sensor for High-Speed Imaging. International Symposium on Solid-state Circuit, pp. 104-105, 2000.

[7] W. Kosonocky, G. Yang, R. Kabra, C. Ye, Z. Pektas, J. Lowrance, V. Mastrocola, F. Shallcross, and V. Patel. 360×360 Element Three-Phrase Very High Frame Rate Burst Image Sensor: Design, Operation, and Performance. IEEE Trans. On Electron Devices, vol. 44, no. 10, pp.1617-1624, 1997.

[8] X. Arreguit, F. Schaik, F. Bauduin, M. Bidiville, and E. Raeber. A CMOS Motion Detector System for Pointing Devices. IEEE Journal of Solid-State Circuits, vol. 31, no. 12, pp. 1916-1921, 1996.

[9] R. Etienne-Cummings, J. Spiegel, and P. Mueller. A Focal Plane Visual Motion Measurement Sensor. IEEE Trans. On Circuits and Systems I: Fundamental Theory and Applications, vol. 44, no. 1, pp.55-65, 1997

[10] R. Harrison, C. Koch. A Robust Analog VLSI Reichardt Motion Sensor. Analog Integrated Circuits and Signal Processing, vol.24, no. 3, pp. 213-229, 2000.

[11] C. Higgins, C. Koch. A Modular Multi-chip Neuromorphic Architecture for Real-Time Visual Motion Processing. Analog Integrated Circuits and Signal Processing, vol.24, no. 3, pp. 195-211, 2000.

[12] Z. Lu. Visual Tracking with Subpixel Resolution using an Analog VLSI Computational Sensor. Proceeding of the 2000 IEEE International Symposium on Robotics & Automation, San Francisco, CA. pp. 1676-1681, April 2000.

[13] S. Kavadias, B. Dierickx, D. Scheffer, A. Alaerts, D. Uwaerts, and J. Bogaerts. A Logarithmic Response CMOS Image Sensor with On-Chip Calibration. IEEE Journal of Solid State Circuits, vol. 35, no. 8, pp. 1146-1152, 2000.

[14] S. Decker, R. McGrath, K. Brehmer, and C. Sodini. A 256×256 CMOS Imaging Array with Wide Dynamic Range Pixels and Column-Parallel Digital Output. IEEE Journal of Solid State Circuits, vol. 33, no. 12, pp. 2081-2090, 1998.

[15] D. Yang, A. Gamal, B. Fowler, and H. Tian. A 640×512 CMOS Image Sensor with Ultrawide Dynamic Range Floating-Point Pixel-Level ADC. . IEEE Journal of Solid State Circuits, vol. 34, no. 12, pp. 1821-1834, 1999.

[16] W. Yang. A wide-Dynamic-Range Low-Power Photosensor Array. International Symposium on Solid-state Circuits, pp. 230-231, 1994.

[17] E. Culurciello, R. Etienne-Cummings, K. Boahen. Arbitrated Address Event Representation Digital Image Sensor. International Symposium on Solid-state Circuits, pp. 92-93, 2000.

[18] L. McIlrath. A Low-Power Low-Noise Ultrawide-Dynamic-Range CMOS Imager with Pixel-Parallel A/D Conversion. . IEEE Journal of Solid State Circuits, vol. 36, no. 5, pp. 846-853, 2001.

[19] X. Guo, M. Erwin and J. Harris. Ultra-Wide Dynamic CMOS Imager Using Pixel-Threshold Firing. In Proceedings of 5th World Multiconference on Systemics, Cybernetics and Informatics, vol. xv, pp. 485-489, 2001.

[20] R. Kishore. Time Domain Quantization CMOS Image Sensor System Design and Architecture. Master Thesis, University of Florida, 2001.

[21] Y. Ni, F. Devos, M. Boujrad, and J. Guan. Histogram-Equalization-Based Adaptive Image Sensor for Real-Time Vision. IEEE Journal of Solid-State Circuits, vol. 32, no. 7, pp. 1027-1036, 1997.

[22] V. Brajovic, T. Kanade. A Sorting Image Sensor: An Example of Massively parallel Intensity-to-Time Processing for Low-Latency Computational Sensors. Proceedings of the 1996 IEEE International Symposium on Robotics and Automation, Minneapolis, Minnesota, pp. 1638-1643, April 1996.

[23] A. Iwata, M. Nagata, N. Takeda. Pulse Modulation Circuit Architecture and its Application to Functional Image Sensor. International Symposium on Circuit and System, vol. 2, pp. 301-304, May 2000.

[24] Z. Zhou, B. Pain, E. Fossum. Frame-Transfer CMOS Active Pixel Sensor with Pixel Binning. IEEE Trans. on Electron Devices, vol. 44, no. 10, pp. 1764 –1768, 1997.

[25] S. Kawahito, M. Yoshida, M. Sasaki, K. Umehara, D. Miyazaki,Y. Tadokoro, K. Murata,S. Doushou,and A. Matsuzawa. A CMOS Image Sensor with Analog Two-Dimensional DCT-Based Compression Circuits for One-Chip Cameras. IEEE Journal of Solid-State Circuits, vol. 32, no. 12, pp. 2030-2040, 1997.

[26] S. Kemeny, R. Panicacci, B. Pain, L. Matthies,and E. Fossum. Multiresolution Image Sensor. IEEE Trans. On Circuits and Systems for Video Technology, vo. 7, no. 4, pp. 575-582, 1997.

[27] K. Aizawa, T. Hamamoto, Y. Egi, M. Hatori,and H. Maruyama. On Sensor image compression for High Pixel Rate Imaging: Pixel Parallel and Column Parallel Architectures. International Symposium on Circuit and System, vol. 4, pp. 179-181, 1996.

[28] K. Aizawa, Y. Egi, T. Hamamoto, M. Hatori,M. Abe, H. Maruyama, and H. Otake. Computational Image Sensor for On Sensor Compression. IEEE Trans. On Electron Devices, vol. 44, no. 10, pp. 1724-1730, 1997.

[29] K. Aizawa, H. Ohno, Y. Egi, T. Hamamoto, M. Hatori, H. Maruyama, and J. Yamazaki. On Sensor Image Compression. IEEE Trans. On Circuits and Systems for Video Technology, vol 7, no. 3, pp. 543-548, 1997.

[30] Q. Luo, J. Harris. A Novel Integration of Wavelet Transform for CMOS Imager. 2002 IEEE International Symposium on Circuits and Systems, Vol.3, pp. 325 - 328, 2002.

[31] M. Schanz, W. Brockherde, R. Hauschild, B. Hosticka, and M. Schwarz. Smart CMOS Image Sensor Arrays. IEEE Trans. On Electron Device, vol. 44, no. 10, pp. 1699-1705, 1997.

[32] C. Aw, B. Woolry. A 128×128 Pixel Standard-CMOS Image Sensor with Electronic shutter. IEEE Journal of Solid-State Circuits, vol. 31, no. 12, pp. 1922-1930, 1996.

[33] Y. Iida, E. Oba, K. Mabuchi, H. Nakamura, H. Miura. A ¼ Inch 330k Square Pixel Progressive Scan CMOS Active Pixel Image Sensor. IEEE Journal of Solid-State Circuits, vol. 32, no. 12, pp. 2042-2047, 1997.

[34] T. Blalock, N. Gaddis, K. Nishimura, and T. Knotts. True Color 1024×768 Microdisplay with Analog In-Pixel Pulsewidth Modulation and Retinal Averaging Offset Correction. IEEE Journal of Solid-State Circuits, vol. 36, no. 5, pp. 838-845, 2001.

[35] B. Dierickx, G. Meynants, and N. Scheffer. Near-100% Fill Factor Standard CMOS Active Pixel. IEEE CCD & AIS Workshop, Brugge, Belgium, pp.1, 1997.

[36] T. Lule, M. Wagner, M. Verhoeven, H. Keller, and M. Bohm. 1000000-Pixel 120dB Imager in TFA Technology. IEEE Journal of Solid-State Circuits, vol. 35, no. 5, pp. 732-739, 2000.

[37] S. Benthien, T. Lule, B. Schneider, M. Wagner, M. Verhoeven, and M. Bohm. Vertically Integrated Sensors for Advanced Imaging Applications. IEEE Journal of Solid-State Circuits, vol. 35, no. 7, pp. 939-945, 2000.

[38] C. Koch. Seeing chips: Analog VLSI Circuits for Computer Vision. Neural Computation, vol. 1, pp. 184-200, 1989.

[39] A. Moini. Vision Chips. Kluwer Academic Publisher, March 1997.

[40] E. Funatsu, Y. Nitta, Y. Miyake, T. Toyoda, J. Ohta, and K. Kyuma. An Artificial Retina Chip with Current-Mode Focal Plane Image Processing Functions. IEEE Trans on Electron Devices, vol.44, no. 10, pp. 1777-1782, 1997.

[41] M. Schwarz, R. Hauschild, B. Hosticka, J. Huppertz, T. Kneip, S. Kolnsberg, L. Ewe, and H. Trieu. Single-Chip CMOS Image Sensors for a Retina Implant System. IEEE Trans on Circuits and Systems II: Analog and Digital Signal Processing, vol. 46, no. 7, pp. 870-877, 1999.

[42] W. Fang. A Smart Vision System-on-A-Chip Design Based on Programmable Neural Processor Integrated with Active Pixel Sensor. International Symposium on Circuit and System, vol. 2, pp. 128-131, 2000.

[43] L. Mcllrath, V. Clark, P. Duane, R. McGrath, and W. Waskurak. Design and Analysis of a 512×768 Current-Mediated Active Pixel Array Image Sensor. IEEE Trans. On Electron Devices, vol. 44, no. 10, pp. 1706-1715, 1997.

[44] J. Nakamura, B. Pain, T. Nomoto, T. Nakamura, and E. Fossum. On-Focal-Plane Signal Processing for Current-Mode Active Pixel Sensors. IEEE Trans. On Electron Devices, vol. 44, no. 10, pp. 1747-1758, 1997.

[45] J. Coulombe, M. Sawan, C. Wang. Variable Resolution CMOS Current Mode Active Pixel Sensor. International Symposium on Circuit and System, vol. 2, pp. 293-296, 2000.

[46] S. Donati. Photodetectors – Devices, Circuits, and Applications. Prentice Hall PTR, 2000.

[47] M. Zambuto. Semiconductor Devices. McGraw-Hill Book Company, 1989.

[48] A. Gamal. EE392B: Introduction to Image Sensors and Digital Cameras. http://www-isl.stanford.edu/~abbas/group, spring quarter 2000.

[49] P. Gray, R. Meyer. Analog Integrated Circuits (3rd Edition). Wiley, 1992.

[50] M. Erwin. Measurements of Test Pixels with $N^+$-diff, n-well, and $P^+$-diff Photodiodes. University of Florida Individual Work Report, supervisor Dr. John G. Harris, May 2001.

[51] H. Zimmermann. Integrated Silicon Optoelectronics. Springer, 2000.

[52] D. Dragoman and M. Dragoman. Advanced Optoelectronic Devices. Springer, 1998.

[53] A. Astrom, R. Forchheimer and P. Danielsson. Intensity Mapping Within the Context of Near-Sensor Image Processing. IEEE Trans. On Image Processing, vol. 7, no. 12, pp. 1736-1741, 1998.

[54] D. Stoppa, A. Simoni, L. Gonzo, M. Gottardi and G. Betta. A 138dB Dynamic Range CMOS Image Sensor with New Pixel Architecture. 2002 IEEE International Solid State Circuits Conference, pp40-41.

[55] P. Debevec. Recovering High Dynamic Range Radiance Maps from Photographs. http://www.debevec.org/Research/HDR, 2002

[56] G. Larson. Overcoming Gamut and Dynamic Range Limitations in Digital Images. In processing of Color Imaging Conference: Color Science, Systems and Applications, Scottsdale, AZ, pp214-219, Nov. 1998.

[57] R. VanRullen and S. Thorpe. Rate Coding versus temporal Order Coding: What the Retinal Ganglion Cells Tell the Visual Cortex. Neural Computation, vol. 13. no. 6, pp. 1255-1283, 2001.

[58] R. Gonzalez and R. Woods. Digital Image Processing. Addison-Wesley Publishing Company, 1992.

[59] A. Yukawa. A CMOS 8-bit High-Speed A/D Converter IC. IEEE J. Solid State Circuits, vol. SC-20, no. 3, pp. 775-779, June 1985.

[60] M. Deen, J. Wang, Z. X and Z. Zuo. Substrate Bias Effects on short channel Length and Narrow Channel Width PMOS devices at Cryogenic Temperatures. Proc. of Workshop on Low Temperature Semiconductor Electronics, pp. 53-56, 1989.

[61] B. Razavi, B. Wooley. Design Techniques for High-Speed, High-Resolution Comparators. IEEE J. Solid State Circuits, vol. 27, no. 12, pp. 1916-1926, 1992.

[62] D. Johns, K. Martin. Analog Integrated Circuit Design. John Wiley & Sons, Inc., 1995.

[63] A. Roychaudhuri. Substrate Bias Dependence of Short-Channel MOSFET Threshold Voltage – A Novel Approach. IEEE Trans. Electron Devices, vol. 35, no. 2, pp.167-173, 1988.

[64] C. Fayomi, G. Roberts and M. Sawan. Low Power/Low Voltage High Speed CMOS Differential Track and Latch Comparator with Rail-to-Rail Input. International Symposium on Circuit and System, vol. 5, pp. 653-656, 2000.

[65] H. Tian, B. Fowler and A. Gamal. Analysis of Temporal Noise in CMOS Photodiode Active Pixel Sensor. IEEE Journal of Solid-State Circuits, vol. 36, no. 1, pp. 92-101, 2001.

[66] A. Ziel. Noise in Solid State Devices and Circuits. John Wiley & Sons, Inc., 1986.

[67] X. Qi, Q. Luo, X. Guo and J. Harris. SNR and FPN Considerations for Time-Based wide Dynamic Range CMOS Image Sensors. In Proceedings of 6th World Multiconference on Systemics, Cybernetics and Informatics, vol. xx, pp. 66-73, 2002

[68] A. Gamal. Introduction to Image Sensors and Digital Cameras. http://www.stanford.edu/class/ee392b, 2000.

[69] B. Dierickx and G. Meynants. Missing Pixel Correction Algorithm for Image Sensors. Processing of SPIE, vol. 3410, pp. 200-203, 1998.

[70] G. Meynants and B. Dierickx. A circuit for the Correction of Pixel Defects in Image Sensors. Processing of Europe Solid-state Circuit, pp. 312-316, 1998.

[71] M. Nagata, J. Nagai, K. Hijikata, T. Morie and A. Iwata. Physical Design Guides for Substrate Noise Reduction in CMOS Digital Circuits. IEEE journal of Solid-State Circuits, vol. 36, no. 3, pp. 539-549, 2001.

[72] T. Lule, S. Benthien, H. Keller, F. Mutze, P. Rieve, K. Seibel, M. Sommer and M. Bohm. Sensitivity of CMOS Based Imagers and Scaling Perspectives. IEEE Trans. on Electron Devices, vol. 47, no. 11, pp. 2110-2122, 2000.

# BIOGRAPHICAL SKETCH

Qiang Luo was born in Nanchang, Jiangxi province, China, in 1973. He completed his bachelor's degree in electrical engineering from Fudan University in July 1995 with honor. He part-time worked as an undergraduate advisor in the electrical engineering department of Fudan University from 1995 to 1997 and completed his master's degree in electrical engineering from Fudan University in July 1998. He began his Ph.D. study at the University of Florida in August 1998 and started the research of the ultra-wide dynamic range CMOS imager from 2000 under Dr. John G. Harris. He was an intern engineer in Texas Instruments Inc. (Dallas, TX) from May 2001 to August 2001, working on the development of an ultra-wide dynamic range CMOS image sensor under Dr. Zhiliang Julian Chen. He received his Ph.D. degree in electrical engineering from the University of Florida in December 2002.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

John G. Harris, Chairman
Associate Professor of Electrical and
Computer Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Robert M. Fox
Associate Professor of Electrical and
Computer Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Jose C. Principe
Distinguished Professor of Electrical and
Computer Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Murali Rao
Professor of Mathematics

This dissertation was submitted to the Graduate Faculty of the College of Engineering and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

December 2002

Pramod Khargonekar

Pramod P. Khargonekar
Dean, College of Engineering

Winfred M. Phillips
Dean, Graduate School